

71

Are small investors naive about incentives?

Ulrike Malmendier ^{a*}, Devin Shanthikumar ^b

^a *University of California, Department of Economics, 549 Evans Hall, #3880,
Berkeley, CA, 94720-3880, USA*

^b *Harvard Business School, Soldiers Field Road, Morgan Hall 377,
Boston, MA, 02163, USA*

***Journal of Financial Economics*, August 2007, vol. 85 (2), pp. 457-489.**

Abstract

Security analysts tend to bias stock recommendations upward, particularly if they are affiliated with the underwriter. We analyze how investors account for such distortions. Using the NYSE Trades and Quotations database, we find that large traders adjust their trading response downward: they exert buy pressure following strong buy recommendations, no reaction to buy recommendations, and selling pressure following hold recommendations. This “discounting” is even more pronounced when the analyst has an underwriter affiliation. Small traders, instead, follow recommendations literally. They exert positive pressure following both buy and strong buy recommendations and zero pressure following hold recommendations. We discuss possible explanations for the differences in trading response, including information costs and investor naiveté.

JEL Classifications: G14, G25, G29, D82, D83.

Keywords: Stock recommendations; Trade reaction; Individual and institutional investors; Conflicts of interest; Behavioral finance

We would like to thank Nick Barberis, Stefano DellaVigna, Ming Huang, Ilan Kremer, Charles Lee, Roni Michaely, Marco Ottaviani, Oguz Ozbas, Josh Pollet, Paul Schultz, René Stulz, Adam Szeidl, Richard Thaler, Kent Womack, an anonymous referee, and seminar participants at Harvard, London Business School, Northwestern University, Stanford, Università Bocconi (IGIER), University of Florida, University of Illinois at Urbana-Champaign, University of Madison-Wisconsin, USC, UT Austin, University of Utah, and Washington University in St. Louis as well as the SITE (Economics & Psychology) 2003, NBER Behavioral Finance Fall 2003, WFA 2004, N.Y.Fed/Ohio State University/JFE 2004, and HBS IMO 2005 conferences for very helpful comments. Michael Jung provided excellent research assistance.

*Corresponding author contact information: ulrike@econ.berkeley.edu

1. Introduction

Stock recommendations of security analysts exhibit a strong upward bias. While the scale of recommendations ranges from “strong sell” and “sell” to “hold,” “buy,” and “strong buy,” only 4.5% of all recommendations recorded in the IBES data set through December 2002 are in the strong sell and sell categories. Analysts’ true scale appears to be shifted upward. The upward bias is even more pronounced for analysts who are affiliated with the underwriter of the recommended stock.

In this paper, we document the trade reaction of investors to recommendations. Using the NYSE Trades and Quotations database, we investigate how large and small traders respond to recommendations issued by affiliated and unaffiliated analysts.

We find three main results. First, both large and small traders display significant trade reactions. But only large traders adjust their trading response to the upward distortion. They exhibit a positive abnormal trade reaction to strong buy recommendations, *no* reaction to buy recommendations, and significant *selling* pressure after holds. Small traders, by contrast, follow recommendations literally. They exhibit a positive abnormal reaction to both buy and strong buy recommendations and no reaction to holds. Second, large traders react significantly less positively to buy and strong buy recommendations if the analyst is affiliated (their overall reaction is insignificantly negative). Small traders, instead, do not respond differently to affiliated recommendations. Third, small investors appear to take less account of the informational content of a recommendation change (or the lack thereof). For example, small investors respond positively to mere reiterations of buy and strong buy (unaffiliated) recommendations, while large investors do not display any significant reaction. The results are robust to alternative econometric specifications,

including alternative investor and analyst classifications, controls for analyst and brokerage heterogeneity, and tests for front running of large traders.

Our results reveal systematic and robust differences in how small and large investors react to analyst reports. It is harder to pin down the explanation for those differences. One possibility is that information about analyst distortions is more costly for small investors – the costs of adjusting their trading behavior outweigh the benefits. In fact, the benefits could be small or even zero due to the arbitrage of large investors. Alternatively, small investors might not seek (or internalize) information about analyst distortions even if the costs of obtaining such information are low. They take recommendations at face value and trust analysts too much, in line with experimental results on advice-giving and the literature on investors' reaction to firms' accounting choices and security issuance decisions (Schotter, 2003; Daniel, Hirshleifer, and Teoh, 2002).

To differentiate between these explanations would require estimates of the costs of and returns to information about analyst distortions. However, informational costs are hard to measure objectively. The returns are, in principle, easier to calculate, but the NYSE Trades and Quotations database does not allow such calculations since it does not reveal investors' portfolio strategies (only aggregate trade imbalances).

As a second-best approach, we analyze the relation between abnormal returns and trade imbalance. Using an event-study methodology, we find that small investors' net (buy minus sell) trade reaction predicts significantly lower abnormal returns than large investors' net trade reaction over six and twelve months. The difference is insignificant if we assume a three-month holding period. We also calculate the portfolio returns to a trading strategy that takes recommendations literally, i.e., buys after buy and strong buy recommendations and sells after sell and strong sell recommendations. Using the Fama-French four-factor portfolio method, we find mostly insignificant abnormal returns.

Two additional results shed some light on the underlying motives of small investors. First, investors face 94.5% positive and neutral recommendations, revealing the *general* distortion at no (additional) cost to those who trade in response to recommendations. Thus, rational small investors should be aware of the *general* upward shift of recommendations by *all* analysts. Nevertheless, they fail to account for it. Second, while it might be costly to distinguish affiliated and unaffiliated analysts and to identify the *additional* distortion of affiliated analysts, small investors can minimize the cost by focusing on analysts who are most easily identified as “independent”: analysts whose financial institutions are *never* involved in underwriting. However, we find that small investors display *less* abnormal trade reaction to such analyst recommendations.

Our paper builds on a large literature on the informational distortions of analysts (Francis, Hanna, and Philbrick, 1997; Lin, McNichols, and O’Brien, 2003, among many others). Several papers document that the recommendations of affiliated analysts are more favorable than those of unaffiliated analysts (see, e.g., Dugar and Nathan, 1995; Lin and McNichols, 1998; and Michaely and Womack, 1999). The high ratio of buy over sell recommendations indicates that even unaffiliated analysts do not provide a balanced view (Michaely and Womack, 2005).¹

Previous analyses of investor reaction to recommendations have been largely based on return patterns. Womack (1996) finds significant three-day event returns to recommendation changes in the direction of the change. The evidence on return differences if analysts are affiliated is mixed. For initial public offering (IPO) underwriting affiliation, Michaely and Womack (1999) show that both the initial positive reaction to upgrades and the post-recommendation drift are stronger if the analyst is unaffiliated. For secondary equity offering (SEO) underwriting affiliation, Lin

¹ Optimism in forecasts, price targets, and long-term growth forecasts, even among unaffiliated analysts, point in the same direction. See, for example, Rajan and Servaes (1997), Dechow, Hutton, and Sloan (1999), Chan, Karceski, and Lakonishok (2002), Brav, Lehavy, and Michaely (2003), and Brav and Lehavy (2003). Malmendier and Shanthikumar (2004) suggest, however, that distortion in recommendations does not necessarily correlate with distortion in earnings forecasts.

and McNichols (1998) find that the market reacts significantly more negatively to affiliated than to unaffiliated hold recommendations, but they do not find significant differences in longer run. Iskoz (2002) shows that institutions account for analyst bias, as far as one can deduce from quarterly institutional ownership data. Mikhail et al. (2006) also analyze the reaction of small and large investors to recommendations, but use dollar trading volume. Their general results are consistent with our findings, though they do not find significant results for affiliated recommendations, possibly due to the skewness of the dollar measure for large trades.

We complement the previous findings in three ways. First, we document the trading response to affiliated and unaffiliated recommendations using measures of buyer and seller initiation as in Odders-White (2000). Second, we distinguish between small and large investors, using the trade-size algorithm developed in Lee and Radhakrishna (2000). We show that large investors — a group dominated by firms and their associated professionals — account for analyst distortions, but small investors do not. Third, we investigate the costs of and returns to adjusting for analyst distortions and relate them to different explanations for the observed trade reaction.

In the remainder of the paper, we first provide details on the various data sources (Section 2), including the classification of investor types and evidence of analyst distortions. Section 3 presents our core result, documenting the trade reaction of small and large investors to analyst recommendations. In Section 4, we discuss several potential explanations and provide a partial analysis of the costs of and the returns to informed trading. Section 5 concludes.

2. Data

2.1. Data sources

We examine data on securities trading, analyst recommendations, and underwriting. The raw trading data are from the New York Stock Exchange Trades and Quotations database (TAQ), which reports every quote and round-lot trade since January 1, 1993 on NYSE, AMEX, and

NASDAQ. We examine ordinary common shares of U.S. firms traded on NYSE. We exclude AMEX and NASDAQ data both since the Lee-Ready algorithm, used to measure trade responses, is error-prone if multiple market makers produce quotes and since the trade-size investor classification is based on NYSE data (Lee and Radhakrishna, 2000; Odders-White, 2000).

We obtain sell-side analyst recommendations and brokerage information since October 29, 1993 from IBES. IBES converts recommendations into a uniform numerical format. We reverse the original IBES coding to 5 = strong buy, 4 = buy, 3 = hold, 2 = sell, and 1 = strong sell. Thus, an “upgrade” translates into a positive change in the numerical value.

We identify upgrades, downgrades, and reiterations relative to the previous recommendation on the same stock by the same brokerage. An initiation is the first recommendation of a brokerage for a stock or, if the brokerage had previously stopped coverage of the stock, a new recommendation. As a robustness check, we use analyst identifiers (rather than brokerage firm identifiers) to classify recommendations. Both methods are largely identical because brokerage firms generally only have one outstanding recommendation on a stock. The analyst-based method forgoes “anonymous” recommendations (with default analyst code 0).

In order to account for left-censoring of the data, which prevents the classification of recommendations at the beginning of our sample period, we drop the first 179 days of the IBES sample period (corresponding to the median time between recommendation updates) when splitting the sample into initiations and other types of recommendations. Any recommendation after 179 days (April 26, 1994) with no preceding recommendation for the same stock by the same brokerage is classified as an initiation. Alternatively, we drop only those recommendations within the first 179 days that cannot be classified, giving more weight to recommendations that are updated more frequently. All results (see the lower half of Table IV) remain virtually identical.

The IBES data contain an unusually high number of recommendations during the first three months, raising concerns about the consistency of the early data.² To account for these reporting anomalies and also to exclude the “scandal effects” of 2001 and 2002 as well as the effects of NASD Rule 2711 on the distribution of recommendations (Barber, Lehavy, McNichols, and Trueman, 2004), we focus on the period February 1994 through July 2001, containing 2,252 securities and 2,229 firms, but we have checked the robustness of the results to using the entire IBES sample period (October 29, 1993 through December 31, 2002).

We classify analysts (or brokerages) as “affiliated” if they belong to a bank that has an underwriting relationship with the firms they are reporting on. As in previous literature (Lin and McNichols, 1998; Michaely and Womack, 1999), we require that the bank was the lead underwriter in an IPO in the past five years or an SEO in the past two years, or a co-underwriter over the same periods. We also examine two sources of underwriting bias that have not been explored previously: SEO underwriting in the next one or two years, and lead underwriting of bonds in the past year. Future underwriters might issue higher recommendations to gain business, to increase future offer prices, or due to winner’s curse. For bond underwriters, positive coverage could be part of an implicit agreement with the issuer, as it is for equity issues. Our analysis focuses on the traditional affiliation measures, both to conform to previous literature and to minimize informational asymmetries between large and small investors, e.g., about future underwriting.

We obtain underwriting data for 1987-2002 from the Securities Data Corporation (SDC) New Issues database and link it to IBES brokerage firms by company name (from the IBES recom-

² While the number of recommendations per year (and per month) is fairly uniform from 2/1994-12/2001, the first two months and three days contain substantially more observations. This could reflect large layoffs in the securities industry: The number of analysts and stocks covered declines sharply, from 626 analysts and 1,166 stocks in 11/1993 to 435 analysts and 591 stocks in 2/1994. However, monthly data from U.S. Dept. of Labor Statistics (DOLS) indicate that the drop in employment is not as sharp as the IBES data suggest. That may be because the DOLS data includes all employees in the securities industry, and equity analysts might have been laid off disproportionately. But it also leaves room for concerns about data consistency within the IBES sample.

mendation broker identification file). We improve the match using company websites and news articles, in particular to determine subsidiary relationships and corporate name changes, and using the mapping of Kolasinski and Kothari (2004).³

We obtain security prices, returns, and shares outstanding from the Center for Research in Security Prices (CRSP), and company variables from COMPUSTAT. The merged data set extends from October 29, 1993 to December 31, 2002 (with underwriting data since 1987), and contains 173,950 recommendations with linked trading data, for 2,424 securities of 2,397 firms. Only 12% of firms lack recommendations, so that our final sample contains almost the entire set of domestic NYSE firms with common stock. We refine the return data by setting returns equal to zero in cases where CRSP codes returns as missing because they are missing on a given day or the last valid price is more than ten days old. From a holding-period perspective, the effective returns are zero in both cases.

2.2. Investor classification

We consider separately the trading behavior of large and small investors. Large traders are likely to be institutional investors, such as pension funds; small traders are more likely to be individual investors. While the composition of the two groups of traders is not crucial for our analysis, it suggests a number of reasons why large traders, but not small traders adjust for analyst distortions, as documented in the literature Behavioral Industrial Organization (Ellison, 2006; DellaVigna and Malmendier, 2004). First, professional investment managers spend their full working time on investment decisions. Repetition, more frequent feedback, and specialization facilitate learning about analyst incentives. Second, finance professionals have a better financial education and better investment skills than the average individual, as illustrated by the

³ We are grateful to Adam Kolasinski and S.P. Kothari for providing us with their mapping, which refines the matches using corporate websites, LexisNexis, Hoover's Online, and the Directory of Corporate Affiliations.

anomalous trade reaction of small traders to earnings news (see, e.g., Lee, 1992; Bhattacharya, 2001; and Shanthikumar, 2003). Finally, when individuals follow bad investment recommendations they forgo returns but will continue to manage their personal funds. Institutions, instead, lose investors and are driven out of the market. Though institutions might not invest optimally, due to misaligned incentives and managerial entrenchment (Lakonishok, Shleifer, and Vishny, 1992), they are more likely to overcome the informational distortions in recommendations.

We distinguish small and large investors by trade size. Lee and Radhakrishna (2000) propose dollar cutoffs based on the three-month Transaction Order and Quote (TORQ) sample from 1990-1991, which reveals the identity of traders and thus allows verifying the accuracy of the classification as individual or institutional investors. Lee and Radhakrishna remove medium-size trades to minimize noise. We follow their suggestion and use \$20,000 and \$50,000 cutoffs. Our results are robust to variations in cutoffs and using share-based rather than dollar-based cutoffs.

We take additional steps to examine the reliability of the trade-size classification, especially given recent changes in trade sizes documented by Kaniel, Saar, and Titman (2005). We obtain data on the portfolio size of individual investors from 1992 until 1998 from the Federal Reserve Bank's *Survey of Consumer Finances* and until 2002 from the *Equity Ownership in America* study of the Securities Industry Association and the Investment Company Institute. In each year, we find that 60-75% of the portfolios of individuals are smaller than \$50,000. Thus, individual trades above \$50,000 are unlikely. A third source, the *NYSE Fact Book*, documents trade sizes directly. From 1991 to 2001, the categories of small (up to 2,099 shares), medium (2,100-4,999 shares), and large (5,000 shares and above) trades had very stable market shares of 20%, 10%, and 70% respectively. After 2001, the shares of large and small trades converged to about 45% each. The stability until 2001 confirms that the Lee-Radhakrishna investor classification applies to our sample period — but changes shortly afterwards.

As a last empirical check, we analyze a large non-public dataset of individual accounts at a large discount brokerage firm over the period 1991-1996.⁴ The vast majority of individual trading lies below \$20,000. The mean (median) trade size is \$12,300 (\$5,256) for common stock. The 90th percentile lies below \$30,000, and even the 95th percentile is below the \$50,000 cutoff. Thus, the data corroborates our categorization of small traders for the subset of retail investors.

We conclude that the Lee-Radhakrishna classification is likely to perform properly in the 1990s and worse thereafter.⁵ To test for time trends, we repeat our core analysis year by year.

2.3. Distortions of analyst recommendations

Sell-side analysts face a well-known conflict of interest when providing investment advice. On the one hand, reliable recommendations attract customers and enhance the analyst's reputation. On the other hand, buy recommendations are more likely to generate trading business than sell recommendations, given short-selling constraints. Moreover, management tends to complain about low ratings and to "freeze out" the issuing analysts, and buy-side clients push for positive recommendations on stocks that they hold.⁶ Analysts face additional pressures if their brokerage is affiliated. Favorable recommendations are generally viewed as an implicit condition of underwriting contracts (Michaely and Womack, 1999; Lin, McNichols, and O'Brien, 2003; Bradley, Jordan, and Ritter, 2003; Conrad, Cornell, Landsman, and Rountree, 2004). Directly or indirectly, analyst compensation depends on the "support" in generating corporate finance profits (Hong and Kubik, 2003; Chan et al., 2003). Sorting can enhance the distortion. Analysts might

⁴ We thank Terry Odean and Itamar Simonson for the data.

⁵ Increasing internalization and trade-shredding are among the reasons for the changes after the 1990s. See, for example, *Wall Street & Technology*, "The Market Makers' Makeover – Decimal pricing and razor-thin profit margins are pushing wholesale market makers to overhaul their trading operations," 7/1/2003, and the *Wall Street Journal*, "SEC Urges U.S. Stock Markets To Help Stop Splitting of Trades," 1/25/2005.

⁶ See Lin and McNichols (1998); Francis, Hanna, and Philbrick (1997); and International Organization of Securities Commissions (2003). Boni and Womack (2002) cite several press reports and the testimony of the (then) acting SEC chairman Laura Unger to the House Subcommittee on July 31, 2001.

choose to cover companies they judge favorably, hoping that those are of most interest to their clients. If they do not account for winner's curse, their recommendations will be too positive.

Previous literature has shown that analyst recommendations are, indeed, systematically shifted upward. We confirm this pattern in our data. Table I displays the sample statistics of affiliated and unaffiliated recommendations for the entire sample period (October 1993 through December 2002), containing 121,130 recommendations. There are 8,466 (7%) affiliated recommendations. Affiliated analysts issue more positive recommendations than unaffiliated analysts. The average recommendation level for any type of affiliated analyst lies around 4, i.e., is at least a "buy." For unaffiliated analysts, the average is statistically significantly lower, at 3.76. Likewise, the mode is "buy" for affiliated analysts but "hold" for unaffiliated analysts.

Analysts make very few sell and strong sell recommendations (4.58%), regardless of their affiliation, but affiliated analysts make even fewer. For example, analysts with IPO and SEO lead and co-underwriter affiliations issue a total of only 154 sell and strong sell recommendations. Affiliated recommendations are significantly higher than unaffiliated ones over the entire sample, but the difference is even stronger before 2002. These differences do not arise from quicker reaction to news. As shown in Malmendier and Shanthikumar (2004), affiliated analysts update their recommendations more slowly (every 357 days) than unaffiliated analysts (every 308 days). While affiliated analysts update negative and hold recommendations faster, they preserve positive recommendations about 70 days longer than unaffiliated analysts.

We also consider separately independent brokerage firms, which do not underwrite any securities, starting five years before our sample period (1987-2002). Such "never-affiliated" brokerages have no corporate finance department and are easily identifiable, even by small investors. Using this definition, 5.3% of the recommendations (6,418) in our sample are independent. Independent analysts make the most sell and strong sell recommendations. Their average recommen-

ation is significantly lower than for unaffiliated analysts (by -0.0622 , with a standard error of 0.0119) or any other group of affiliated analysts (the difference to all affiliated analysts is -0.3102 , with a standard error of 0.0157). As a robustness check, we classify analysts as independent if they have no underwriting affiliation during the five years prior to a recommendation, which applies to 15.3% of recommendations (18,486). This definition has the advantage of including information only up to the recommendation date and the disadvantages of failing to capture “true independence” (the additional 10% have a corporate finance department) and being harder to identify by small investors. We replicate all results using the alternative measure.

Finally, we test whether the more positive recommendations of affiliated analysts reflect differences in the firms they cover. Companies that have recently issued securities might be better investments, as evidenced by their access to capital markets. In Panel B of Table I, we restrict the sample to firms that have recently issued stocks or bonds. The statistics are virtually identical. In addition, a detailed comparison of the distribution of covered stocks across the National Association of Investors Corporation (NAIC) industries reveals very close similarities. The portion of affiliated and unaffiliated recommendations falling into any of the NAIC industry groups differs by less than one percentage point for all but three industries.

In summary, recommendations display two types of bias. First, more than 95% are positive or neutral. Second, recommendations are even more positive if the analyst is affiliated.

3. Trade reaction

How do investors react to recommendations? We distinguish between large and small investors and ask how they account for the two types of distortions.

3.1. Measuring trade reaction

We measure the trade reaction to recommendations with the directional trade-initiation measures of Lee and Ready (1991). “Buyer (seller) initiated” means that the buyer (seller) demands

immediate execution, generally representing a market order, which is to be executed immediately at the current market price. Thus trade initiations capture buy and sell pressure.

We use the modified version of the Lee-Ready algorithm developed by Odders-White (2000). The algorithm matches a trade to the most recent quote that precedes the trade by at least five seconds. If a price is nearer the ask price it is classified as buyer-initiated; if it is nearer the bid price, it is classified as seller-initiated. If a trade is at the midpoint of the bid-ask spread, it is classified based on a “tick test.” The tick test categorizes a trade as buyer-initiated (seller-initiated) if the trade occurs at a price higher (lower) than the price of the previous trade, i.e., on an uptick (downtick). We drop trades at the bid-ask midpoint, which is also the same price as in preceding trades.

As a proxy for net buy pressure, we consider three measures. The net number of buyer-initiated trades for firm i , investor type x , and date t is defined as

$$NB_{i,x,t} = buys_{i,x,t} - sells_{i,x,t} \quad (1)$$

The raw trade imbalance measure for firm i , investor type x , and date t is calculated as

$$TI_{i,x,t} = \frac{buys_{i,x,t} - sells_{i,x,t}}{buys_{i,x,t} + sells_{i,x,t}} \quad (2)$$

We normalize this measure by subtracting the firm-year mean and dividing by the firm-year standard deviation, separately for each investor type, as in Shanthikumar (2003):

$$TI_{i,x,t}^{abnormal} = \frac{TI_{i,x,t} - \overline{TI}_{i,x,year(t)}}{SD(TI_{i,x,year(t)})} \quad (3)$$

The adjustments are made by year to account for changes in trading behavior over time and by firm to account for differences in small and large trading behavior for different stocks. These normalizations allow us to compare trading behavior over time and among firms and replace year- and firm-fixed effects in the regression framework.

As a robustness check, we also replicate our analysis using the number of shares or the dollar amount of trades rather than the number of trades.

3.1.1. Ownership

Trade initiation is different from “making a trade.” It captures how urgently investors want to trade. While investors place limit orders for reasons such as liquidity trading, trade initiation indicates that an investor has a strong belief about future stock price movements.

The algorithm for trade initiation does not identify the other side of the trade. A buyer-initiated large trade, for example, could be filled with a large non-initiated sell order, with several small trades that are pulled together, or by the market maker. Thus, trade imbalances do not *necessarily* lead to ownership changes between different investor groups.

In order to test whether or not trade imbalances predict ownership changes in our data, we compare them to changes in institutional ownership in the CDA/Spectrum Institutional Holdings data (13f SEC filings). Since the data are quarterly, we aggregate the trade measures over the corresponding quarters. Table II displays correlations with ownership change. We find that large-trader buy pressure is significantly correlated with an increase in institutional ownership, and small-trader buy pressure with a decrease. The results are even stronger if we adjust for time trends in institutional ownership by removing the average ownership change in a given quarter. We also find consistent results when measuring the correlations for each quarter *separately*. Despite the loss of power, the correlation of the sum of daily trade imbalances and institutional ownership is significant at the 10% level in 21 of the 36 quarters, in the expected directions.

To gauge the economic significance of these correlations, we regress the quarterly changes in institutional ownership on the quarterly sum of large-investor trade imbalances and a constant. We find that an increase of one standard deviation in the summed large trade imbalances more than doubles the average increase in institutional ownership, from 0.4% to 0.86%. Thus, trade

imbalance corresponds to statistically and economically significant changes in ownership composition, as far as we can infer from the quarterly institutional ownership data.

3.2. Trade reaction of small and large investors

To analyze trade reactions to recommendations, we distinguish trading on the first two days at or after recommendation issuance (event days 0 and 1) from the remaining sample period.

Table III presents summary statistics. Panel A shows that, over the full sample period, small investors execute 8.49 more buyer-initiated trades and 8.73 more seller-initiated trades per day than large investors. The average differences between buyer- and seller-initiated trades are very similar, 3.18 for small trades and 3.43 for large trades. The median difference is zero for both small and large trades. During the two event days, the differences between buys and sells are considerably higher, 19.26 for small trades and 18.92 for large trades (Panel B). Thus, recommendations induce systematic buy-pressure among all investors — a first indication that they have informational value even for large investors, consistent with Barber et al. (2001) and Jegadeesh et al. (2004). The normalized trade imbalance for large traders is slightly negative (un-normalized it is slightly positive, 0.058), indicating that large portion of trades initiated by large traders on event days are for firms with high average trade imbalances.

In Table IV, we regress the normalized abnormal trade reaction on dummies for each recommendation level and interactions with an affiliation dummy, separately for large and small traders. The coefficients of the five level effects show the reaction to unaffiliated recommendations, and the five interactions show the differential reaction if the analyst is affiliated. In a third column, we report the difference between large and small investors' trade reaction. All standard errors are adjusted for heteroskedasticity and arbitrary within-event-date correlation.

The first four panels (upper half of Table IV) show the regression results for the full set of all recommendations during our main sample period (February 1994 through July 2001) and three

subsamples: 1) excluding three-day windows around earnings announcements; 2) excluding reiterations; and 3) excluding both earnings announcements and reiterations. Subsample 1 allows us to distinguish the effect of recommendations from the effect of simultaneous earnings announcements. We consider a plus-or-minus one-day window around the announcement date, since the reaction of small traders is strongest on days 0 and 1 after the announcement (Battalio and Mendenhall, 2005; Shanthikumar, 2003). Using the IBES data on earnings announcements to identify the overlapping dates,⁷ we find that 12% of recommendations occur during the three trading days around earnings announcements, similar to the joint distribution of recommendations and earnings announcement in Womack (1996). Subsample 2 separates out reiterations, which should have little or no informational content and thus no impact on trades. When identifying reiterations we account for the left-censoring of the data by dropping the first 179 days of the IBES sample period (i.e., before April 26, 1994), as described in Section 2.1. Subsample 3 combines the restrictions of Subsamples 1 and 2.

In the full sample (first three columns), large investors' reaction to unaffiliated recommendations is significantly positive for strong buy recommendations, insignificantly positive for buy recommendations, and significantly negative for hold, sell, and strong sell recommendations. Small investors, instead, display significantly positive reaction to both buy and strong buy recommendations and zero trade reaction to hold recommendations. They display negative abnormal trading responses only to sell and strong sell recommendations.

The implications of these baseline results are two-fold. First, recommendations have a significant impact on the trading behavior of *both* large and small investors. Second, large traders ac-

⁷ The quality of matches between recommendations and earnings announcements is high: 99.35% of the recommendations have a matched earnings announcement date within ± 80 days. Alternatively, we used the COMPUSTAT data to identify earnings announcement dates. The resulting fraction of matches within the same ± 80 -day window is lower (93.69%), plausibly due to the lower-quality matching mechanism (six-digit CUSIP rather

count for the upward bias in analyst recommendations by shifting their reaction down by one level (e.g., they hold in response to buy recommendations and sell in response to hold recommendations). Small traders do not adjust their trades but take analyst recommendations literally. The upper part of Fig. 1 (“Unaffiliated Analysts”) summarizes schematically the differences in behavior between traders who adjust for the upward distortion (Column 1) and traders who do not adjust (Column 2).

The interaction coefficients in the next five rows show that large investors react significantly less positively to strong buy or buy recommendations if the analyst is affiliated. Their overall reaction is zero (insignificantly negative). The differential reaction to hold, sell, and strong sell recommendations is insignificant. Thus, as with unaffiliated recommendations, they react negatively to neutral and negative recommendations. Small investors, instead, do not display any significant difference in their reaction to positive recommendations if the analyst is affiliated. They react more negatively to affiliated strong sell recommendations.

The results indicate that large traders apply an additional downward adjustment to positive recommendations if the analyst is affiliated. Small investors, instead, do not differentiate between affiliated and unaffiliated recommendations. They take both types of recommendations literally. The lower part of Fig. 1 (“Affiliated Analysts”) summarizes these differences in behavior, displaying the net trade reaction (rather than the differential as in Table IV).

Note that the schematic reaction “without adjustment” (Column 2 of Fig. 1) does not predict that the coefficient of small investors’ trade reaction is more negative for affiliated than for unaf-

than IBES ticker). We also repeated our analysis for the subsample of recommendations that have a matched (IBES) earnings announcement dates. The results are virtually identical.

filiated strong sells (-0.838 versus -0.105 in Table IV). This result is due to an IBES coding imprecision⁸ and is identified out of an extremely small sample (27).

As in all empirical work on trade reactions, the coefficient of determination is rather low, around 1%, revealing large cross-sectional heterogeneity. Since the focus of the analysis is not to forecast trade volume but to contrast small and large investors' reaction to recommendations, the goodness of fit has a very limited role.

The normalization of our trade imbalance measure also allows us to compare the magnitude across investor groups. The coefficients indicate that small traders react more strongly than large traders to positive recommendations. Their reaction to unaffiliated strong buy recommendations is about two times as large as that of large investors, and their reaction to buy recommendations is about ten times as large. The reactions to negative recommendations are similar in magnitude for both types of investors (except the affiliated strong sell anomaly discussed above). The weaker reaction of large investors could reflect earlier access to information, including front running as discussed below, or additional discounting for the upward distortions.

It is harder to interpret the economic significance of the coefficients. Since the standard deviation of abnormal trade imbalance is normalized to 1, we can interpret the coefficients*100 as a percentage of the standard deviation. For example, an unaffiliated strong buy recommendation triggers an increase of 11% of one standard deviation of abnormal trade imbalance among large traders, but 24% among small traders.

The next three panels (next nine columns in the upper half of Table IV) reveal that the results are not driven by earnings announcements or reiterations. After removing recommendations

⁸ Some analysts use four instead of five recommendation levels, only one of which is negative. IBES codes the negative level as *strong sell*. Four categories turn out to be particularly common among affiliated analysts: 30% of the 27 affiliated recommendations coded as strong sell represent a general *sell* category. This affects the dynamics of the negative trade reaction. Since analysts most often downgrade in stages, strong sells are typically preceded by

within three days around earnings announcements for the same stock, we estimate very similar coefficients. All coefficients follow the exact same pattern of being significantly positive, significantly negative, or insignificant. The coefficients of the large hold-affiliation interaction and of the small hold dummy switch sign but remain insignificant. For example, while large traders display a significantly negative reaction to unaffiliated holds both in the full sample and in the sample without earnings announcement days (-0.091 with a standard error of 0.011, and -0.095 with a standard error of 0.012, respectively), the small traders' reaction changes from small and insignificantly positive (0.007, standard error 0.014) to small and insignificantly negative (-0.015, standard error 0.014). In both samples, the small-trader coefficient is much smaller in absolute magnitude than the large-trader coefficient. And both differences between the large-trader reaction and the small-trader reaction are highly significant.

The same is true for Subsample 2. After removing reiterations, all 20 estimated coefficients are again of the same direction (or remain insignificant) and of very similar magnitude. Large traders' reaction to buy recommendations is now closer to significant (0.018, standard error 0.012) though still an order of magnitude smaller than the small-trader reaction (0.141, standard error 0.014). In Subsample 3, where we remove both recommendations around earnings announcements and reiterations, all 20 coefficients have again the same pattern of significantly positive, significantly negative, and insignificant. As with Subsamples 1 and 2, a few coefficients switch sign (large-trader hold and small-trader hold, hold interaction, and strong-buy interaction); but as before, these coefficients are insignificant and small in absolute magnitude. Thus, neither simultaneous information in earnings announcements nor uninformative reiterations appear to be driving our results.

3.2.1. Initiations, reiterations, upgrades, and downgrades

sells. Small investors who normally sell "already" in response to the earlier sells rather than the (subsequent) strong

Prior recommendations affect the informational content of recommendations. For example, upgrades to buy convey more positive information than buy reiterations. To test which types of recommendations are driving our full-sample results, we split the sample into initiations, reiterations, upgrades, and downgrades. As with Subsample 3, we remove the first 179 days of the IBES sample and exclude recommendations within three days around same-stock earnings announcements. The majority of recommendations (43%) are initiations. Only 5% are reiterations. The remainder is roughly equally split between upgrades (24%) and downgrades (28%). The relative portions of initiations, upgrades, and downgrades are similar to previous studies, e.g., Barber, Lehavy, McNichols, and Trueman (2001). The percentage of reiterations is lower in IBES than in previously used data.

The lower half of Table IV shows the regression results for each subsample. The categories of affiliated sell and strong sell reiterations and affiliated sell upgrades disappear due to lack of data. The category of affiliated sell initiations remains but contains only one recommendation.

We find that the positive reaction to strong buy recommendations, which we estimated for both large and small investors in the full sample, is fairly consistent among all three relevant subsamples. The exception is large traders' reaction to strong buy reiterations, which is insignificantly negative. Small traders, instead, display a significantly positive trade reaction to these reiterations. The coefficients for unaffiliated buy recommendations, for which we estimated a small (insignificantly positive) reaction among large traders and a large (significantly positive) reaction among small traders in the full sample, are more divergent across subsamples. For small traders, the coefficient estimate is positive in every subsample, though insignificant for downgrades. For large traders, we estimate a negative coefficient in the case of reiterations (insignificant and small) and downgrades (significant and large). The initiations estimate is again similar

sells are now coded as reacting to the "strong sells," triggering the higher abnormal sell reaction.

to the full sample (small, positive, and insignificant), and the upgrades estimate is significantly positive. Overall, as shown in columns “Difference S-L,” large traders react significantly less positively to buy and strong buy recommendations than small traders in every subsample.

For unaffiliated holds, we find that the negative reaction of large traders, which we estimated on the full sample, reflects a large negative response to downgrades (-0.162) and initiations (-0.033). The response is insignificant for reiterations and upgrades. Small traders’ reaction to holds is instead small and insignificant in every subsample.

Unaffiliated sell and strong sell recommendations, for which we estimated negative trading responses for both types of investors, trigger the strongest negative response in the downgrades sample. Most other coefficient estimates are also negative though typically insignificant.

Turning to the differential reaction to affiliation, we find that the discounting of large traders after positive recommendations replicates in the initiations sample. In the other subsamples (and for other recommendation levels), the interaction coefficient is insignificant, with two exceptions: the coefficient on affiliated sell initiations is large and significantly negative, though identified out of one observation; the coefficient on affiliated upgrades to hold is also significantly negative. For small traders we estimate mostly insignificant differential trading responses, with four exceptions: the differential reaction to sell initiations is significantly negative, that to affiliated strong buy reiterations significantly negative, and that to buy and strong buy upgrades significantly positive. The first result is estimated out of a single observation. The others remain puzzling. They reveal, however, that upgrades are driving the “affiliation neglect” of small traders, while initiations are driving the “affiliation adjustment” of large traders.

In summary, we can link the full-sample estimates to three main subsample results. First, the key results replicate in the initiations subsample. Large traders react less positively than small traders to unaffiliated hold, buy, and strong buy recommendations, and they discount positive

recommendations even more if the analyst is affiliated. Sell and strong sell initiations, however, trigger insignificant responses by both large and small traders. The less negative and less significant reaction relates to the results in Bradley, Jordan, and Ritter (2003): initiations right after the quiet period trigger less response; later initiations trigger a particularly positive response.

Second, reiterations cause no statistically significant trading response among large traders. For large traders, the negative coefficient of the hold-affiliation interaction, -0.288, is closest to significant with a t-statistic of 1.43. All others are even farther from significant, with t-statistics between 0.15 and 0.88. By contrast, reiterations trigger a significantly positive reaction among small traders in the case of unaffiliated buy and strong buy recommendations. Thus, small traders appear to account less for the informational content of recommendation changes.

Third, for a given recommendation level, large traders react more positively to upgrades and more negatively to downgrades. Small traders' reaction is less closely aligned with the direction of changes, e.g., when displaying an insignificantly positive (rather than significantly negative) reaction to downgrades from strong buy to buy. Small traders *do* react significantly more positively to upgrades than to downgrades in the case of buys (t-statistic 4.64) but not in the case of hold recommendations (t-statistic 1.43).⁹ As mentioned above, the results for strong sell recommendations are more mixed, possibly due to the small sample sizes. (Only 2.6% of the 1,172 strong sells are reiterations; 23.6% are initiations; the remainder are downgrades.)

3.3. Robustness

We perform several robustness checks of the documented trading behavior.

3.3.1. Econometric model

⁹ Note that the negative sell-coefficients in the upgrades, reiterations, and initiations sample are larger than in the downgrades sample but insignificant for upgrades and reiterations. The estimation is affected by the extremely small sample size of the identifying subsample. There are only 64 upgrades to sell and 49 reiterations of sell recommendations, while there are 1,283 downgrades to sell and 398 initiations at the sell level.

Panel A of Table V reestimates the standard errors, allowing for heteroskedasticity and arbitrary within-year correlation (columns marked “Cluster by Year”) and arbitrary within-brokerage firm correlation (columns marked “Cluster by Brokerage Firm”). Clustering by year allows for a wider range of cross-correlations than our primary method (clustering by date), although the low number of clusters is problematic (Wooldridge, 2002; Froot, 1989). Our results are robust to these alternative assumptions. Large investors adjust downwards and display a sell reaction to hold and no reaction to buy recommendations. They also react more negatively to affiliated than to unaffiliated strong buy recommendations, although the downward adjustment is not significant for buy recommendations. The abnormal trade reaction of small investors follows again a literal (unadjusted) interpretation of recommendations and does not differentiate between affiliated and unaffiliated recommendations (as before with the exception of strong sells).

We also repeat the regressions of Table IV including year- and brokerage-fixed effects. Alternatively, we include year-firm interactions, adding more than 20,000 fixed-effects groups, in order to capture firm characteristics that change over time. As expected, given that the measure of trade imbalance is normalized on a firm-year basis, the results do not change.

3.3.2. Investor type classification

We check the robustness to several variations in the cutoff values for trade size (Table V.B): \$1-\$5,000, \$5,000-\$10,000, \$10,000-\$20,000, and \$20,000-\$50,000. Our baseline small-investor cutoff aggregates the first three groups. Both sets of results for small traders — the literal reaction to recommendations and the lack of adjustment for affiliation — replicate in almost all cells. The largest group (\$20,000-\$50,000) behaves more like large investors. Traders discount positive recommendations if the analysts are affiliated. The interaction coefficients for buy and strong buy are significantly negative. Moreover, the puzzling differential trade reaction to affiliated strong sells, which we found in Table IV, loses significance in the largest group. On the

other hand, the reaction to negative unaffiliated recommendations becomes insignificant in two subgroups (below \$5,000 and \$20,000-\$50,000). Overall, both results show remarkable robustness within each of these small subgroups.

3.3.3. Measure of trade reaction

The results are similar if we employ the net number of buyer- minus seller-initiated trades or the raw trade imbalance. Also, using normalized imbalances of shares traded or of dollar amounts traded (rather than trades) produces similar results: 19 of the 20 coefficients and all ten differences between small and large traders display the exact same pattern of significantly positive, significantly negative, or insignificant. Only the large-trader interaction of affiliation and buy recommendations changes, to insignificantly negative.

Longer horizons (up to 20 trading days after the recommendation) also lead to similar results. Finally, we limit the sample to firms that have at least *some* institutional ownership at the quarter-end before the recommendation, using several different cutoffs, as large trading is most likely to be from institutions for these firms, and find similar results.

3.3.4. Affiliation

We also perform several robustness tests of our affiliation classification. First, we split “affiliation” into its component parts, and we use the additional forms of affiliation, future underwriting and bond underwriting. Second, we specify whether a firm has recently issued a security and whether the underwriter is independent. Third, we repeat the baseline regression separately for each year in the sample. With all of these variations, the primary results remain.

3.3.5. Relationship between investor type and affiliation

Another concern is that investors systematically lower their trading size in response to affiliation. The resulting (re-)classification of large investors as small could generate their weaker reaction to affiliated buy recommendations.

An immediate weakness of this alternative explanation is that it cannot explain our first result, i.e., that small investors do not discount analyst recommendations on average. Two additional results address the concern directly. First, systematic shifts in trade size that are large enough to move investors normally trading above \$50,000 into the class of trades below \$20,000 should be reflected in the remaining class of large trades. For example, a uniform shift would reduce the average size of the large trades by at least \$30,000. However, the average size of large trades is \$217,244 in response to unaffiliated recommendations and \$209,836 in response to affiliated recommendations, a reduction of less than 3.5%.

Second, in order to explain the more negative reaction of large traders in response to affiliated recommendations, a general reduction of trade size does not suffice. Rather, the proportion of buy-initiators among large traders has to go down. Instead, both changes are small and similar. Buyer-initiated large trades change on average by 3.9% and seller-initiated large trades by 2.5%.

3.3.6. *Analyst heterogeneity*

The differences in trading responses could arise if small investors follow different (more affiliated) analysts than large investors. To address this concern, we reduce the heterogeneity of recommendations and consider only those issued by brokerages that are both affiliated and unaffiliated at different points during the sample period (“Ever-Affiliated Brokerages,” Columns 1 to 3 in Table VI). All results replicate, for both large and small traders.

As a second way to address heterogeneity, we restrict the analysis to analysts who were listed in *Institutional Investor*’s most recent October list of top analysts (“All Star Team,” Columns 4 to 6).¹⁰ The resulting sample is significantly smaller (11,882 observations), but most results replicate. Exceptions are the differential response of large investors to affiliated buys (now insignificant) and of small investors to affiliated strong buys (now marginally significant with $t =$

1.821 in the All-Star sample). The reaction to negative recommendations also loses significance in several instances. Overall, the core results replicate in both subsets and analyst heterogeneity and adverse sorting of small investors appear unlikely to generate our results.

3.3.7. Brokerage firm heterogeneity

Even if small investors do not follow worse analysts, they may only be aware of more widely known brokerage firms. To address this concern, we obtain sales and employee data for brokerages from the Dun & Bradstreet Million Dollar Database, which is available for about 5% of our sample. We control for brokerage size by interacting every level of recommendation with sales or with the number of employees as proxies for size. In unreported regressions, we find that the inclusion of size controls does not diminish the results beyond the effect of the reduced sample.

3.3.8. Front running of large investors

Another important concern is that large investors might trade prior to recommendations. If they learn the information that sparks new recommendations earlier, their trade reaction could take place earlier. At the time of the recommendation, they then display either no reaction or a contrarian reaction, explaining their lower event-time trade reaction.

In order to test for front running, we calculate the average daily trading volume of large investors during the month prior to the recommendation. All averages—number of trades, shares traded, and dollar volume traded—peak on the day of the recommendation and not before. We also observe a slow increase over time. This trend could reflect anticipatory trading before the release of the recommendation or a response to general news about the stock. To distinguish these explanations, we re-analyze the relation between recommendation and trade imbalance, as in Table IV, for the week *preceding* recommendations (days -5 to -1). None of the coefficients implies anticipatory trades among large traders. Instead, large traders exhibit a significant buy-imbalance

¹⁰ We thank Steven Drucker for providing us with the 1995-2001 lists of “All Star” analysts. We obtained the

for “downgrade buys,” i.e., stocks that are currently strong buys but will be downgraded to buy within the next five trading days, while they exhibit a significant sell-imbalance once the downgrade occurs. Similarly, they exhibit insignificant negative pressure before an upgrade to strong buy, but significant buying pressure when the upgrade occurs. These results remain unchanged when we include a larger event window.

4. Interpretation

What explains small investors’ lack of adjustment to 1) the *general* upward bias and 2) the *specific* upward bias of affiliated analysts? One potential reason is different information costs. If it is less costly for institutional investors than for individuals to find out about analyst distortions, only institutions might choose to acquire the information. An alternative explanation, which we dub investor naiveté, is that some small investors fail to adjust their trades even to freely available information about analyst distortion, or that they underinvest in information acquisition.

Evaluating the different explanations requires measures of the cost of information and the return implications of the trading responses. If costs are zero or low enough that it would increase small trader’s utility if they acquired the information, there is room for investor naiveté. Otherwise, we can explain their behavior in the standard rational framework.

Such an evaluation is difficult to perform. First, it is hard to measure the costs of obtaining information. Second, measuring the returns to information requires information about the portfolios of small investors, which is not contained in our data. Third, even if we could measure the returns, they do not easily translate into utility. Nevertheless, a few additional results shed some light on the costs and benefits of information and thus the plausibility of the explanations.

4.1. Benefits of information about analyst distortion

names for the remainder of our sample period using the October issues of *Institutional Investor Magazine*.

A plausible proxy for the benefits of adjusting to analyst distortion is the difference in returns earned by small and large investors' trade reactions, neglecting their (possibly heterogeneous) translation into utility. However, our data do not reveal portfolio choices. For example, we do not observe the typical holding period in response to different types of recommendations. As a second best, we use trade imbalances to approximate the trading strategies in two ways.

4.1.1. Trade reaction and event returns

First, we test whether the direction and strength of buy or sell pressure among each investor group predicts event returns. We use buy-and-hold returns net of the value-weighted CRSP market returns. The market-adjusted return of stock j on day t is

$$A_{jt} = R_{jt} - R_{mt}$$

We regress the abnormal return on a constant and on the dollar value of net buyer- minus seller-initiated trades on event days 0 and 1. We perform this analysis over three, six, and twelve months after each recommendation. As shown in Table VII, Panel A, abnormal trades by small investors predict significantly negative returns over the six-month horizon and insignificantly negative returns over three and twelve months (with p-values of 13% and 10%, respectively). Large traders' trade reaction predicts instead significantly positive abnormal returns over all horizons. The difference between the coefficients for large and small traders is significant for six and twelve months and insignificant (at a p-value of 12%) for three months. Thus, if we assume holding periods of six or twelve months, small traders incur losses relative to large traders from their reactions to recommendations. The estimated loss is not significant for three months.

4.1.2. Analyst recommendations and portfolio returns

Our second approach is to examine the returns to investment strategies that follow analyst recommendations "literally": if an analyst issues a buy or strong buy recommendation, the investor purchases the stock in the long portfolio; if the analyst issues a sell or strong sell recommenda-

tion, the investor sells (shorts) the stock in the short portfolio. We consider holding periods of three, six, and twelve months. A stock is dropped from the portfolio when the analyst revises the recommendation to any level other than hold, stops covering the stock, or when the holding period expires. If an analyst issues a hold recommendation during the holding period of a stock, the holding period restarts. If multiple analysts make the same type of recommendation for the same stock, the stock appears in the portfolio multiple times. The portfolio is reevaluated daily. Changes occur at the end of the trading day of the corresponding recommendation event. We split the analysis into affiliated and unaffiliated recommendations.¹¹ Whenever the short portfolio is empty, we drop it from our calculations. We use the Fama-French four-factor portfolio method to determine value-weighted buy-and-hold abnormal returns separately for the long and the short portfolio, estimating

$$R_{it} - R_{ft} = \alpha_i + \beta_i (R_{mt} - R_{ft}) + s_i SMB_t + h_i HML_t + u_i UMD_t + \varepsilon_{it} \quad (4)$$

where $R_{i,t}$ is the return of portfolio i , $i \in \{Buy, Sell\}$, on day t ; $R_{m,t}$ is the return of the market portfolio on day t ; $R_{f,t}$ is the risk-free rate on day t ; SMB , HML , and UMD are the size, book-to-market and momentum factors;¹² and α_i estimates the (gross) abnormal return. For the short portfolio, we multiply all coefficients by -1 to display the returns to selling rather than buying.

We calculate the net daily abnormal returns by subtracting the estimated brokerage commission and bid-ask spread. We apply the \$30 (online) per-trade commission at Charles Schwab during our sample period to the average trade size of small investors, \$8,971.73. We add the 1% cost of the bid-ask spread estimate of Barber and Odean (2000) for individual investors, and es-

¹¹ The unaffiliated long portfolio contains, on average, 1,571, 2,835, and 4,617 stocks for the holding periods of three, six, and twelve months; the affiliated long portfolio contains only 107, 197, and 328 stocks on average for the same horizons. The affiliated long portfolio is missing for the first day of our sample period, since no affiliated recommendation was issued. Similarly, the unaffiliated short portfolio consists of 103, 184, and 285 stocks, on average, while the affiliated short portfolio is empty for 201 days and contains, on average, only two stocks during the remaining 1,690 days, for the three-month horizon; it is empty during 24 days, with an average of four stocks, for the six-month horizon, and it is empty for twelve days, with an average of five stocks, for the twelve-month horizon.

timate total transaction costs of 1.669%. (We neglect the market impact of trading, which would further increase the transaction cost.) We apply the transaction cost estimate to the daily turnover of each portfolio.¹³ Note that the volume of trading, while not excessive, likely exceeds real-world trading of most small investors. However, since we neglect the market impact of trading and since we use the size of *actual* trades in our sample to estimate trading commissions, our transaction cost estimate is conservative relative to the, say, 2.99% roundtrip commission in Barber and Odean (2000), restricting the transaction cost estimate. We also recalculate returns for a similar strategy with only monthly rebalancing, with even stronger results (see below).

Panel B of Table VII presents the daily returns (in %) of the resulting portfolios. In the long portfolio, gross abnormal returns are insignificantly positive for unaffiliated recommendations and mostly negative and insignificant for affiliated recommendations. Accounting for transaction costs, the returns are always negative and, in three out of six cases, strongly significant, with t-statistics between 1.94 and 3.25 (otherwise between 0.18 and 1.49). In each case, the abnormal returns are worse in the portfolio of affiliated analysts than the portfolio of unaffiliated analysts. In the short portfolio, *selling* stocks with negative recommendations induces negative abnormal returns both gross and net of transaction costs in each of the six portfolios, though significantly so only in three cases after transaction costs. Here as well, the estimates are always more negative for affiliated than for unaffiliated analysts, but the differences are not significant.

All results are similar and more significant when using later transaction dates, e.g., the end of the first day or the end of the second day after the recommendation. The abnormal returns deviate at most by 0.008%, except for the affiliated short portfolio, which deviates by up to 0.014%.

¹² *SML*, *HML*, $R_{m,t}$ and $R_{f,t}$ are from Ken French's website; daily *UMD* factors are from Jeffrey Busse.

¹³ To calculate turnover, we construct a comparison portfolio with no rebalancing from the previous day; the only changes in weight are due to returns. We obtain the difference for each stock between the actual weight and the hypothetical weight without rebalancing and sum the positive differences. By summing only the positive differences, we avoid double-counting and can apply the full round-trip transaction cost directly. The average daily turnover

Gross abnormal returns are lower in most cases, and net abnormal returns are significantly negative at the 1% level in seven of the twelve portfolios.

The results are also stronger if we use monthly instead of daily rebalancing. We consider a strategy of updating only once a month while still taking recommendations “literally.” At the end of each month, the investor trades on recommendations made during that month. Despite the lower turnover, the net returns are still significantly negative. Moreover, the gross return estimates are more consistently negative: the returns from buying any of the three affiliated long portfolios and the returns from selling any of the six (affiliated and unaffiliated) short portfolios are negative, though insignificant. The affiliated portfolio performs always worse than the corresponding unaffiliated portfolio. The difference is significant for the three-month long portfolio.

We also calculate the returns of the zero-investment portfolios constructed from the long and short portfolios in Table VII, Panel B. The abnormal returns are negative over any horizon, insignificantly so before transaction costs, and significantly in all but one case after transaction costs. The differences between returns to the affiliated and the unaffiliated portfolios widen, relative to the separate long and short portfolios, though they remain insignificant. The losses from following affiliated recommendations are sizable, with stable negative abnormal returns of -0.04% to -0.07% before transaction costs and -0.10% to -0.12% after transaction costs. The returns of the unaffiliated zero-investment portfolio amount to one-third to two-thirds of those magnitudes.¹⁴

Finally, we repeat the analysis separately for stocks with low institutional interest. If small investors affect stock prices, we would expect their impact to be larger on such stocks and, thus,

ranges from 0.0064 (for the unaffiliated long portfolio with a one-year holding period) to 0.0270 (for the affiliated long portfolio with a three-month holding period).

¹⁴ The returns of the zero-investment portfolio are to be taken with caution given the small number of stocks in the short portfolio. The average number of “holdings per day” in the long portfolio over those in the short portfolio range from 17:1 to 21:1 for unaffiliated analysts and 67:1 to 100:1 for affiliated analysts.

the returns to their trading strategy to be more negative than for the full universe of stocks. We use two proxies for low institutional interest: 1) stocks in the lowest decile of market capitalization as of the first trading day in a year and 2) stocks covered by only one analyst as of the day of recommendation issuance. The smallest stocks tend to get less institutional interest due to higher transaction costs and regulatory implications of large ownership (e.g., Keim and Madhavan, 1998). Low coverage is a direct proxy for low institutional interest.

Reestimating the model of Eq. (4) separately for each level of recommendation in both subsets of stocks, we find that buy and hold recommendations earn large negative abnormal returns over most horizons. The results are significant even before transaction costs for half of the horizon-affiliation combinations in the case of holds and for a few in the cases of buys.

To sum up, we find significantly negative abnormal portfolio returns to following analyst recommendations literally *without* accounting for transaction costs for stocks with low institutional interest (small and low-coverage stocks). For the overall sample, transaction costs are crucial in turning the mostly negative but insignificant abnormal portfolio returns into significantly negative abnormal returns. Thus, we do not find clear evidence of losses from failing to account for analyst distortions prior to transaction costs.

Our findings confirm, however, the message of Barber and Odean (2000) that “trading is hazardous to [retail investors’] wealth.” One way to interpret our results is that investors would earn significantly higher returns if they forwent trading in response to analyst recommendations. An alternative interpretation is that small investors trade anyhow — with or without recommendations. In other words, recommendations might not be the cause of excess transaction costs.

To evaluate the importance of recommendations for excessive trading, we compare the trade reaction they trigger to the impact of other determinants of trading: same-day abnormal volume and prior-day return, as analyzed in Barber and Odean (2006). Same-day abnormal volume is the

current day's trading volume, normalized by the the stock's prior-year average daily trading volume; prior-day return is the prior trading-day raw return, both as reported in CRSP. For each day, we independently sort stocks into deciles of abnormal volume and prior-day return, and we count the number of recommendations of a given level made for the stock on the same day.

We regress abnormal trade imbalance of small traders, separately, on indicator variables for the decile category of abnormal volume, for the decile category of prior-day return, and on our recommendation count variables. The (unreported) results for abnormal volume and prior-day return are similar to the Barber and Odean results. Abnormal volume is a strong (monotonic) predictor of abnormal buying. Prior-day return is a significant though weaker (and U-shaped) predictor of returns. The highest and lowest trade imbalances due to same-day volume occur in the highest and lowest volume decile and amount to 0.0615 and -0.0780. The highest and lowest trade imbalances due to prior-day returns are 0.0679 (for the lowest returns decile) and -0.0271 (in the fifth decile). In comparison, a single buy recommendation raises the abnormal trade imbalance by 0.0323 and a strong buy recommendation by 0.0613. Limiting the sample to stock-days with recommendations, we find that a single buy recommendation raises the abnormal trade imbalance by 0.0753, and a strong buy recommendation by 0.1052. Even though these magnitudes are not straightforward to compare, recommendations appear to be nearly as strong a predictor of trading as same-day volume and possibly stronger than prior-day returns. Thus, while the possibility remains that small investors would have traded anyway, the comparison to other triggers of abnormal trade suggests that recommendations are likely an important determinant of "trading too much."

4.1.3. Results in prior literature

Our results might at first appear be in contrast with a literature documenting that highly recommended stocks outperform the market (e.g., Womack, 1996; Barber et al., 2001). The differ-

ence, however, is readily explained with the different strategies. The prior papers form buy portfolios using recommendation consensus, which implies strong recommendations with substantial agreement among analysts. In Barber et al. (2001), for example, the most-recommended portfolio contains stocks on which at least half of the covering analysts have a strong buy and the rest a buy recommendation. This portfolio excludes many of the firms in our long portfolio. In addition, “literal” trading uses recommendation levels rather than changes. Boni and Womack (2006) show that portfolios formed from recommendation changes are more profitable.

In sum, the event returns point to negative return implications of small investors’ trade reactions, relative to that of large investors, if we assume six- or twelve-month holding periods. The difference is insignificantly negative over three months. The portfolio returns are insignificant. Thus, a rational model with costly information could suffice to explain small investor behavior. Unfortunately, our return analysis is only suggestive, given the lack of data on trading strategies.

4.2. Cost of information about analyst distortion

We evaluate the cost of information in three steps. First, we ask whether small investors are able to access the full set of recommendations. Second, we ask whether the cost of information can explain the *general* failure to adjust to distortions of *all* analysts. Third, we ask whether the cost of information can explain the difference in small and large traders’ responses to affiliation.

4.2.1. Information about recommendations

Recommendations are published in various forms, such as analyst reports, online sources (such as *briefing.com*, *FirstCall* of Thomson Financial, and *finance.yahoo.com*), radio and TV interviews, and news articles. Online resources provide information on averages, industry comparisons, upgrades, and downgrades. In addition, companies post their press releases (including information about issuances and underwriters) online. Investors can get to the releases with a few clicks from the *Yahoo!* site.

The increasing convenience of online information suggests that recommendations have become more accessible over time. Over our sample period, however, trade reactions to recommendations do not display any time pattern. The results of Table IV are replicable year-by-year, confirming the view of Michaely and Womack (2005) that recommendations became broadly accessible after the 1980s.

The recommendations available to the general public, however, could differ from those available to large investors. To quantify the accessibility to small investors, we collect information from brokerage publications and websites about their target customer: institutions or individuals. We obtain the information for about 85% of the brokerage firms. (A significant number of firms have been acquired or dissolved.) Of those, only 16.2% target institutions (e.g., DSP/Merrill Lynch or SG Cowen).¹⁵ Even those recommendations are often available to retail clients for two reasons. First, numerous institution-oriented firms have acquired or made minority investments in retail brokerage firms (to have a retail distribution outlet for IPO shares), such as Morgan Stanley acquiring Dean Witter and Salomon Brothers acquiring Smith Barney. Second, retail brokerages allow clients to download reports from firms with which they have partnerships (such as Charles Schwab with Goldman Sachs). Individual investors, therefore, have access to most if not all recommendations in our data.

Given that investors can access a large sample of recommendations, informational costs cannot explain our first finding that small investors fail to adjust for *general* upward distortion among *all* analysts. Since 95% of recommendations are positive or neutral, even small investors

¹⁵ We identified, for example, Adams, Harkness & Hill as serving just institutional clients from their mission statement: "Adams, Harkness & Hill is one of the largest independent research, brokerage, and investment banking firms serving the institutional market." We identified Alliance as serving both markets from their statement: "At Alliance Capital, we're proud to provide a wide range of investment management services to a diverse group of investors worldwide, including institutional clients, high-net-worth individuals and mutual fund investors." We also used information about the lines of business and services offered. Research firms with only 10-15 analysts that are very specialized in a particular field, e.g., energy or healthcare, typically serve only selected institutional investors.

should be aware of the general distortion. Instead, small investors do not discount recommendations while large investors do. Thus, our first finding is hard to explain in a rational framework, while consistent with investor naiveté.

4.2.2. Information about affiliation

Small investors could face substantial costs to find out which analysts are affiliated. These costs are a plausible explanation of our second finding that small investors do not differentiate between affiliated and unaffiliated recommendations.

However, if high information costs are the explanation, small investors should look for low-cost ways to obtain this information. Small investors could focus on analysts who are “independent” and never involved in any underwriting. These firms often advertise their independence. In our data, 105 (27%) of the 382 brokerage firms never underwrite during the entire sample period (starting in 1987). They issue 5% of the recommendations. The recommendations of these analysts are less upward-biased than in any other affiliated or unaffiliated subsample (Table I). They also induce less (if any) underperformance, though the differences are not significant. Applying the same portfolio strategy as in Panel B of Table IV, we find that the gross alphas from buying the three long portfolios (over three, six, and twelve months) are up to 0.02% higher per day. Those from selling the short portfolio are up to 0.07% higher.¹⁶

In Table VIII, we test whether small investors attempt to overcome informational constraints by focusing on independent analysts.¹⁷ We find that they react *less* to positive recommendations from independent brokerages, with t-statistics of 1.7 and 1.8 for the buy and the strong buy inter-

¹⁶ The differences in alpha are 0.0099, 0.0212, and 0.0149 relative to the three-, six-, and twelve-month affiliated long portfolio; 0.0336, 0.0731, and 0.0698 for the three-, six-, and twelve-month unaffiliated short portfolio; 0.0009, 0.0013, and -0.0006 for the three-, six-, and twelve-month unaffiliated long portfolio; and 0.0139, 0.0271, and 0.0223 for the three-, six-, and twelve-month unaffiliated short portfolio.

¹⁷ Large investors do not need to focus on independent recommendations since they adjust their trade reaction to the distortions. Nevertheless, we find that large investors do not display any significant downward adjustment to positive recommendations of independent analysts and that they display zero trade reaction (rather than a negative reaction) to independent hold recommendations.

actions, respectively (Column 1). This finding does not appear to be due to the small size of the brokerages. The results also hold for the largest 50% of firms for which we have sales or employee count data (Columns 2 and 3).¹⁸

Overall, informational costs are unlikely to explain why small investors do not account for the *general* distortion of analysts, but are a likely explanation for the undifferentiated reaction to affiliated and unaffiliated recommendations. However, small investors do not appear to choose analyst recommendations in a manner that helps to remedy informational constraints.

5. Conclusion

Analysts tend to positively bias the information they provide to investors, as evident in the very low number of sell and strong sell recommendations, in particular by affiliated analysts. While large investors adjust their reaction to hold and buy recommendations downwards, small investors take recommendations literally. Small investors also fail to account for the additional distortion due to underwriter affiliation. Potential explanations are higher costs of information and naiveté about distortions in analyst recommendations. We shed some light on the plausibility of these explanations by evaluating the costs and benefits of information about analyst distortion. Lacking information about portfolio strategies, we are limited to measuring the event returns to net trade imbalances and the portfolio returns to a strategy that takes analyst recommendations literally. The event returns to small traders' net trade reaction are significantly lower than those of large traders if we assume six- or twelve-month holding periods; the difference is insignificant over three months. The portfolio returns are mostly negative but insignificant, at least before transaction costs. The portfolio returns are significantly negative prior to transaction costs for

¹⁸ The reduced sample size in Columns 2 and 3 induces a few surprising results, such as the insignificantly positive coefficient on strong sell and the significantly positive interaction coefficient of sell and "never affiliated." These results are driven by minimal sample sizes; e.g., there is only one independent sell recommendation. The

stocks with low institutional interest (small and low-coverage stocks). Thus, many of the return results are consistent with a rational model of small investors' trade reaction, though at best a very noisy proxy for the return implications of the response to recommendations.

It is harder to explain in a standard framework why only large traders but not small investors adjust their trade reaction to the *general* upward bias of analyst recommendations, given that there seem to be some conditions under which it would be more profitable to make that adjustment and given that the general upward bias is visible to any trader reacting to recommendations. It is also striking that small traders do not focus on analysts from independent brokerages. The latter findings suggest that small investors are naive about the distortions and trust analysts too much.

small-sample anomalies disappear if we keep the full sample and use a triple interaction of “never affiliated” and “top 50% sales or employees” with the respective recommendation level.

References

- Barber, B., Lehavy, R., McNichols, M., Trueman, B., 2001. Can investors profit from the prophets? Security analyst recommendations and stock returns. *Journal of Finance* 56, 531-563.
- Barber, B., Lehavy, R., McNichols, M., Trueman, B., 2004. Buys, holds, and sells: The distribution of investment banks' stock ratings and the implications for the profitability of analysts' recommendations. Working paper. UC Davis, Davis, CA.
- Barber, B., Odean, T., 2000. Trading is hazardous to your wealth: The common stock investment performance of individual investors. *Journal of Finance* 55, 773-806.
- Barber, B., Odean, T., 2006. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. Working paper. UC Davis, Davis, CA.
- Battalio, R., Mendenhall, R., 2005. Earnings expectations, investor trade size, and anomalous returns around earnings announcements. *Journal of Financial Economics* 77, 289-319.
- Bhattacharya, N., 2001. Investors' trade size and trading responses around earnings announcements: An empirical investigation. *Accounting Review* 76, 221-244.
- Boni, L., Womack, K., 2002. Solving the sell-side research problem: Insights from buy-side professionals. Working paper. Dartmouth College, Hanover, NH.
- Boni, L., Womack, K., 2006. Analysts, industries, and price momentum. *Journal of Financial and Quantitative Analysis* 41, 85-109.
- Bradley, D., Jordan, B., Ritter, J., 2003. The quiet period goes out with a bang. *Journal of Finance* 58, 1-36.
- Brav, A., Lehavy, R., 2003. An empirical analysis of analysts' target prices: Short-term informativeness and long-term dynamics. *Journal of Finance* 58, 1933-1967.
- Brav, A., Lehavy, R., Michaely, R., 2003. Using expectations to test asset pricing models. Working paper. Duke University, Durham, NC.
- Chan, L., Karceski, J., Lakonishok, J., 2003. Analysts' conflict of interest and biases in earnings forecasts. NBER Working Paper No. 9544.
- Conrad, J., Cornell, B., Landsman, W., Rountree, B., 2004. How do analyst recommendations respond to major news? Working paper. University of North Carolina, Chapel Hill, NC.
- Daniel, K., Hirshleifer, D., Teoh, S., 2002. Investor psychology in capital markets: Evidence and policy implications. *Journal of Monetary Economics* 49, 139-209.
- Dechow, P., Hutton, A., Sloan, R., 1999. The relation between analysts' forecasts of long-term earnings growth and stock price performance following equity offerings. *Contemporary Accounting Research* 17, 1-32.
- DellaVigna, S., Malmendier, U., 2004. Contract design and self-control: theory and evidence. *Quarterly Journal of Economics* 119, 353-402.
- Dugar, A., Nathan, S., 1995. The effect of investment banking relationships on financial analysts' earnings forecasts and investment recommendations. *Contemporary Accounting*

Research 12, 131-160.

- Ellison, G., 2006, Bounded rationality in Industrial Organization. Working Paper. Massachusetts Institute of Technology. Cambridge, MA.
- Francis, J., Hanna, J.D., Philbrick, D., 1997. Management communications with securities analysts. *Journal of Accounting and Economics* 24, 363-394.
- Froot, K., 1989. Consistent covariance matrix estimation with cross-sectional dependence and heteroskedasticity in financial data. *Journal of Financial and Quantitative Analysis* 24, 333-355.
- Hanley, K., Lee, C., Seguin, P., 1996. The marketing of closed-end fund IPOs: Evidence from transactions data. *Journal of Financial Intermediation* 5, 127-159.
- Hong, H., Kubik, J., 2003. Analyzing the analysts: Career concerns and biased earnings forecasts. *Journal of Finance* 58, 313-351.
- International Organization of Securities Commissions, 2003. Report of the IOSCO Technical Committee on Analysts Conflicts of Interest (September).
- Iskoz, S., 2002. Relative performance and institutional reaction to underwriter analyst recommendations. Working paper. Massachusetts Institute of Technology, Cambridge, MA.
- Jegadeesh, N., Kim, J., Krische, S., Lee, C., 2004. Analyzing the analysts: When do recommendations add value? *Journal of Finance* 59, 1083-1124.
- Kaniel, R., Saar, G., Titman, S., 2005. Individual investor sentiment and stock returns. Working paper. Duke University, Durham, NC.
- Keim, D., Madhavan, A., 1998. The cost of institutional equity trades. *Financial Analysts Journal* 54, 50-69.
- Kolasinski, A., Kothari, S.P., 2004. Analyst objectivity and investment banking relationships: Evidence on analysts affiliated with M&A advisors. Working paper. Massachusetts Institute of Technology, Cambridge, MA.
- Lakonishok, J., Shleifer, A., Vishny, R., 1992. The structure and performance of the money management industry. *Brookings Papers on Economic Activity: Microeconomics*.
- Lee, C., 1992. Earnings news and small traders: An intraday analysis. *Journal of Accounting and Economics* 15, 265-302.
- Lee, C., Radhakrishna, B., 2000. Inferring investor behavior: Evidence from TORQ data. *Journal of Financial Markets* 3, 83-111.
- Lee, C., Ready, M., 1991. Inferring trade directions from intraday data. *Journal of Finance* 46, 733-746.
- Lin, H., McNichols, M., 1998. Underwriting relationships, analysts' earnings forecasts and investment recommendations. *Journal of Accounting and Economics* 25, 101-127.
- Lin, H., McNichols, M., O'Brien, P., 2003. Analyst impartiality and investment banking relations. Working paper. National Taiwan University, Taipei, Taiwan.

- Malmendier, U., Shanthikumar, D., 2004. Do security analysts speak in two tongues? Working paper. UC Berkeley, Berkeley, CA.
- Michaely, R., Womack, K., 1999. Conflict of interest and the credibility of underwriter analyst recommendations. *Review of Financial Studies* 12, 653-686.
- Michaely, R., Womack, K., 2005. Brokerage recommendations: Stylized characteristics, market responses, and biases. In Thaler, R., (Ed.), *Advances in Behavioral Finance II*, Princeton University Press, Princeton, NJ.
- Mikhail, M., Walther, B., Willis, R., 2006. When security analysts talk, who listens? Working paper. Arizona State University, Phoenix, AZ.
- Odders-White, E., 2000. On the occurrence and consequences of inaccurate trade classification. *Journal of Financial Markets* 3, 259-286.
- Rajan, R., Servaes, H., 1997. Analyst following of initial public offerings. *Journal of Finance* 52, 508-529.
- Schotter, A., 2003. Decision making with naïve advice. *American Economic Review* 93, 196-201.
- Shanthikumar, D., 2003. Small trader reactions to consecutive earnings surprises. Working paper. Harvard University, Boston, MA.
- Womack, K., 1996. Do brokerage analysts' recommendations have investment value? *Journal of Finance* 51, 137-167.
- Wooldridge, J., 2002. *Econometric Analysis of Cross Section and Panel Data*. MIT Press, Cambridge, MA.

TABLE I. Sample of Recommendations

Panel A: Entire Sample	Sample size	Percentage within category						Numerical translation ¹	
		Strong			Strong			Mean	Standard Deviation
		Sell	Sell	Hold	Buy	Buy	Buy		
All	121,130	1.72	2.86	36.84	32.90	25.67	3.78	0.92	
Unaffiliated	112,664	1.79	2.96	37.68	32.40	25.17	3.76	0.92	
Affiliated	8,466	0.73	1.61	25.68	39.56	32.42	4.01	0.84	
IPO lead-underwriting (past 5 years)	1,104	0.63	1.45	23.82	38.41	35.69	4.07	0.84	
SEO lead-underwriting (past 2 years)	1,198	0.42	1.50	21.87	39.90	36.31	4.10	0.82	
Co-underwriting equity ²	4,143	0.99	1.62	26.43	38.79	32.17	4.00	0.86	
Future SEO (next 2 years)	665	0.00	0.30	14.286	45.564	39.85	4.25	0.70	
Bond lead underwriting (past year)	2,083	0.62	1.87	27.99	39.85	29.67	3.96	0.84	
Never Affiliated ³	6,418	3.91	4.25	36.63	28.01	27.19	3.70	1.04	

Panel B: Subsample of firms with an IPO in the past 5 years, an SEO in the past 2 years or a bond issue in the past year	Sample size	Percentage within category						Numerical translation ¹	
		Strong			Strong			Mean	Standard Deviation
		Sell	Sell	Hold	Buy	Buy	Buy		
All	54,952	1.55	2.47	34.99	33.73	27.24	3.83	0.91	
Unaffiliated	46,715	1.70	2.62	36.60	32.73	26.36	3.79	0.92	
Affiliated ⁴	8,237	0.75	1.65	25.88	39.43	32.28	4.01	0.85	

Sample period is 10/29/1993 to 12/31/2002.

¹ The numerical translation scheme is 1=strong sell, 2=sell, 3=hold, 4=buy, 5=strong buy.

² We exclude co-underwriters who are also lead underwriters of SEO or IPO issuances to eliminate the large number of double-counts in this particular category.

³ A brokerage firm is "Never Affiliated" if it does not have any (lead or co-underwriter) equity or bond underwriting affiliation during the entire sample period.

⁴ "Affiliated" summarizes the same categories as in Panel A (IPO in the past 5 years, SEO in the past 2 years, IPO/SEO co-underwriting over the same horizons, future underwriting in the next 2 years, and bond underwriting in the past year).

TABLE II. Correlations of Institutional Ownership and Trades

Correlations of the change in the percentage of shares outstanding owned by institutions (as of 13f SEC filings) and measures of trade reaction. Correlations are calculated using all sample firm-quarters from 12/1993 to 12/2002 with both recommendations data and trading data. P-values in parentheses.

	Small Investors	Large Investors
Sum of daily abnormal trade imbalances over last	-0.073 (0.000)	0.070 (0.000)
Quarterly trade imbalance, number of trades	-0.082 (0.000)	0.088 (0.000)
Quarterly trade imbalance, number of shares	-0.089 (0.000)	0.122 (0.000)
Quarterly trade imbalance, dollar value	-0.085 (0.000)	0.119 (0.000)

TABLE III. Measures of Trade Reaction: Summary Statistics**Panel A. Summary Statistics Daily Trading for Sample Firms (2/1994 - 7/2001)**

	Mean	Median	Std. Dev.	Min	Max
Total number of small trades	66.10	32	102.17	0	3,953
Total number of large trades	45.55	7	128.46	0	3,627
Number of small buy-initiated trades	29.97	13	50.98	0	1,702
Number of large buy-initiated trades	21.49	3	62.14	0	1,911
Number of small sell-initiated trades	26.79	13	42.76	0	2,453
Number of large sell-initiated trades	18.06	3	51.09	0	1,563
Total number of small buy/sell-initiated trades	56.76	26	91.06	0	3,506
Total number of large buy/sell-initiated trades	39.55	6	112.42	0	3,339
Δ (buy-sell) initiated small trades	3.18	0	23.71	-1,440	965
Δ (buy-sell) initiated large trades	3.43	0	17.44	-660	791
Dollar value total small trades	562,641	244,263	932,865	0	27,500,000
Dollar value total large trades	12,200,000	1,082,491	47,200,000	0	6,640,000,000
Dollar value small buy-initiated trades	255,760	99,175	461,493	0	12,300,000
Dollar value large buy-initiated trades	5,579,860	417,750	22,700,000	0	4,860,000,000
Dollar value small sell-initiated trades	228,392	98,550	387,906	0	16,000,000
Dollar value large sell-initiated trades	4,666,593	382,524	18,300,000	0	3,120,000,000
Dollar value total small buy/sell-initiated trades	484,153	204,600	828,517	0	22,700,000
Dollar value total large buy/sell-initiated trades	10,200,000	918,875	40,000,000	0	5,510,000,000
Dollar value of (buy-sell) small trades	27,368	2,338	201,131	-10,600,000	8,854,894
Dollar value of (buy-sell) large trades	913,267	0	9,824,109	-1,430,000,000	4,860,000,000
<i>N</i>	2,996,265				

Panel B. Summary Statistics Trade Imbalance - Sum over Event Days 0 and 1

	Mean	Median	Std. Dev.	Min	Max
Δ (buy-sell) initiated small trades	19.26	5	78.52	-2,545	1,560
Δ (buy-sell) initiated large trades	18.92	3	55.78	-543	1,059
Dollar value of (buy-sell) small trades	161,318	40,428	670,243	-17,100,000	11,800,000
Dollar value of (buy-sell) large trades	4,775,637	369,841	33,900,000	-1,400,000,000	1,520,000,000
Normalized imbalance of small trades	0.1087	0.1265	1.6348	-15.8431	7.1467
Normalized imbalance of large trades	-0.0063	0.0141	1.4083	-9.4254	7.1931
<i>N</i>	86,962				

TABLE IV. Trade Reaction: Regression Results

OLS regressions of normalized trade imbalance over event days 0 and 1 on dummies for recommendation level (Strong Sell, Sell, Hold, Buy, Strong Buy) and affiliation. The sample All Recommendations uses the full I/B/E/S sample over the sample period 2/1994 through 7/2001. The sample All Recommendations Excluding Earnings Announcements drops recommendations that fall within the three-trading-day window around earnings announcements for the same stock. The sample All Recommendations Excluding Reiterations drops reiterations of the same level of recommendation for the same stock by the same brokerage. The last two samples exclude recommendations during the first 179 days of I/B/E/S sample period (median period between recommendation updates) to account for the left censoring of the data. Standard errors (in parentheses) are robust to heteroskedasticity and arbitrary within recommendation-date correlation across firms.

	All Recommendations			All Recommendations Excluding Days of Earnings Announcements			All Recommendations Excluding Reiterations			All Recommendations Excluding Days of Earnings Announcements and Excluding Reiterations		
	Large Trade	Small Trade	Difference S-L	Large Trade	Small Trade	Difference S-L	Large Trade	Small Trade	Difference S-L	Large Trade	Small Trade	Difference S-L
Strong Sell	-0.103 (0.040)	-0.105 (0.050)	-0.002 (0.064)	-0.088 (0.043)	-0.105 (0.053)	-0.018 (0.068)	-0.105 (0.042)	-0.099 (0.051)	0.006 (0.066)	-0.086 (0.045)	-0.102 (0.055)	-0.017 (0.070)
Sell	-0.118 (0.034)	-0.139 (0.046)	-0.021 (0.057)	-0.106 (0.037)	-0.162 (0.048)	-0.056 (0.060)	-0.125 (0.035)	-0.141 (0.047)	-0.017 (0.059)	-0.108 (0.038)	-0.159 (0.049)	-0.051 (0.062)
Hold	-0.091 (0.011)	0.007 (0.014)	0.099 (0.018)	-0.095 (0.012)	-0.015 (0.014)	0.080 (0.019)	-0.097 (0.011)	0.010 (0.014)	0.107 (0.018)	-0.099 (0.012)	-0.013 (0.015)	0.086 (0.019)
Buy	0.011 (0.012)	0.134 (0.013)	0.123 (0.017)	0.007 (0.012)	0.123 (0.014)	0.116 (0.018)	0.018 (0.012)	0.141 (0.014)	0.123 (0.018)	0.014 (0.013)	0.130 (0.014)	0.116 (0.019)
Strong Buy	0.112 (0.013)	0.243 (0.014)	0.131 (0.019)	0.106 (0.013)	0.230 (0.014)	0.124 (0.020)	0.124 (0.013)	0.251 (0.014)	0.127 (0.019)	0.118 (0.013)	0.241 (0.014)	0.123 (0.020)
(Strong Sell)*Affiliation	-0.196 (0.255)	-0.838 (0.331)	-0.643 (0.418)	-0.213 (0.274)	-1.004 (0.348)	-0.791 (0.443)	-0.142 (0.261)	-0.850 (0.344)	-0.709 (0.431)	-0.215 (0.274)	-1.007 (0.348)	-0.792 (0.443)
(Sell)*Affiliation	0.094 (0.254)	-0.087 (0.272)	-0.180 (0.372)	0.152 (0.271)	-0.121 (0.284)	-0.273 (0.393)	0.138 (0.259)	-0.126 (0.277)	-0.264 (0.379)	0.198 (0.278)	-0.172 (0.289)	-0.370 (0.401)
(Hold)*Affiliation	-0.001 (0.044)	0.005 (0.056)	0.006 (0.072)	0.004 (0.049)	0.001 (0.060)	-0.004 (0.077)	0.002 (0.045)	0.000 (0.058)	-0.002 (0.073)	0.007 (0.049)	-0.005 (0.063)	-0.012 (0.080)
(Buy)*Affiliation	-0.068 (0.034)	0.013 (0.039)	0.081 (0.052)	-0.081 (0.037)	0.015 (0.043)	0.096 (0.057)	-0.077 (0.035)	0.004 (0.041)	0.081 (0.054)	-0.087 (0.038)	0.008 (0.044)	0.095 (0.058)
(Strong Buy)*Affiliation	-0.129 (0.036)	-0.023 (0.041)	0.106 (0.055)	-0.126 (0.038)	-0.004 (0.044)	0.122 (0.058)	-0.137 (0.037)	-0.008 (0.044)	0.129 (0.057)	-0.138 (0.039)	0.013 (0.046)	0.151 (0.061)
Sample size	86,961	86,961	86,961	76,457	76,457	76,457	79,831	79,831	79,831	70,138	70,138	70,138
R ²	0.0034	0.0085	0.0085	0.0032	0.0077	0.0077	0.004	0.0091	0.0091	0.0038	0.0085	0.0085

Table IV. Trade Reaction: Regression Results (continued)

OLS regressions of normalized trade imbalance over event days 0 and 1 on dummies for recommendation level (Strong Sell, Sell, Hold, Buy, Strong Buy) and affiliation. Initiations are all first recommendations of a given brokerage for a given stock or recommendations after the brokerage had previously stopped coverage of the stock. Reiterations (Upgrades, Downgrades) are recommendations that are identical to (higher than, lower than) the previous recommendation of a given brokerage for the same stock. All subsample exclude recommendations during the three-day windows surrounding earnings announcements and during first 179 days of the I/B/E/S sample period to account for the left-censoring of the data. Standard errors (in parentheses) are robust to heteroskedasticity and arbitrary within recommendation-date correlation across firms.

	Initiations			Reiterations			Upgrades			Downgrades		
	Large Trade	Small Trade	Difference S-L	Large Trade	Small Trade	Difference S-L	Large Trade	Small Trade	Difference S-L	Large Trade	Small Trade	Difference S-L
Strong Sell	0.012 (0.098)	0.174 (0.109)	0.162 (0.146)	-0.194 (0.273)	-0.325 (0.425)	-0.132 (0.505)	-0.302 (0.175)	-0.301 (0.238)	0.001 (0.295)	-0.119 (0.050)	-0.195 (0.062)	-0.076 (0.079)
Sell	-0.105 (0.074)	-0.126 (0.107)	-0.021 (0.129)	0.043 (0.213)	-0.393 (0.332)	-0.436 (0.395)	-0.302 (0.175)	-0.301 (0.238)	0.001 (0.295)	-0.100 (0.044)	-0.163 (0.055)	-0.063 (0.070)
Hold	-0.033 (0.017)	-0.015 (0.021)	0.019 (0.027)	0.020 (0.057)	0.005 (0.060)	-0.015 (0.082)	-0.014 (0.042)	0.056 (0.049)	0.070 (0.064)	-0.162 (0.015)	-0.019 (0.019)	0.143 (0.024)
Buy	0.028 (0.017)	0.136 (0.020)	0.107 (0.026)	-0.037 (0.041)	0.111 (0.041)	0.148 (0.058)	0.080 (0.019)	0.194 (0.022)	0.114 (0.029)	-0.101 (0.023)	0.035 (0.026)	0.137 (0.035)
Strong Buy	0.095 (0.018)	0.221 (0.019)	0.126 (0.026)	-0.044 (0.056)	0.161 (0.062)	0.205 (0.083)	0.140 (0.018)	0.260 (0.019)	0.120 (0.026)			
(Strong Sell)*Affiliation	-0.863 (0.855)	-0.853 (1.144)	0.009 (1.428)							-0.104 (0.284)	-0.976 (0.359)	-0.872 (0.458)
(Sell)*Affiliation	-1.919 (0.074)	-3.220 (0.107)	-1.302 (0.129)							0.266 (0.277)	-0.060 (0.280)	-0.326 (0.394)
(Hold)*Affiliation	0.102 (0.098)	0.007 (0.117)	-0.095 (0.153)	-0.288 (0.201)	-0.250 (0.296)	0.038 (0.358)	-0.564 (0.254)	0.320 (0.315)	0.883 (0.404)	0.038 (0.058)	-0.016 (0.074)	-0.054 (0.094)
(Buy)*Affiliation	-0.124 (0.051)	-0.085 (0.059)	0.039 (0.078)	0.033 (0.217)	-0.041 (0.215)	-0.074 (0.305)	-0.099 (0.102)	0.265 (0.103)	0.363 (0.145)	0.045 (0.067)	0.118 (0.085)	0.074 (0.109)
(Strong Buy)*Affiliation	-0.200 (0.051)	-0.106 (0.060)	0.095 (0.079)	-0.040 (0.148)	-0.310 (0.165)	-0.271 (0.222)	-0.030 (0.059)	0.204 (0.071)	0.235 (0.092)			
Sample size	31,611	31,611		3,979	3,979		17,780	17,780		20,747	20,747	
R ²	0.0019	0.0081		0.001	0.0052		0.0068	0.0215		0.0102	0.0019	

FIGURE 1. Investor trade reactions to analyst recommendations.

The Figure displays the abnormal trade reaction to analyst recommendations if investors adjust downwards (Column 1) and if they do not adjust (Column 2).

	Recommendations	(1) Downward Adjustment	(2) No Adjustment
Unaffiliated Analysts	strong sell	sell*	(strong) sell*
	sell	sell*	sell
	zero	↓ sell	zero
	buy	↓ weak buy or zero	buy
	strong buy	↓ (less strong) buy	strong buy
Affiliated Analysts	strong sell	sell*	(strong) sell*
	sell	sell*	sell
	zero	↓ sell	zero
	buy	↓ weak buy or zero	buy
	strong buy	↓ (less strong) buy or zero	strong buy

* Alternatively, investors may display no abnormal trade reaction since previous negative (sell or hold) recommendations induced them to sell earlier. Short-sell constraints may strengthen this effect.

TABLE V. Robustness.

Panel A. Clustering

OLS regressions of normalized trade imbalance. Trade reaction is summed over event days 0 and 1. The recommendation levels (Strong Sell, Hold, Buy, Strong Buy) and Affiliation are dummy variables. Cluster by Year and Cluster by Brokerage Firm indicate that the standard errors are robust to heteroskedasticity and arbitrary within-year or, respectively, to arbitrary within-firm correlation. Sample period is 2/1994 through 7/2001. Standard errors in parentheses.

	Cluster by Year			Cluster by Brokerage Firm		
	Large Trade	Small Trade	Difference S-L	Large Trade	Small Trade	Difference S-L
	Strong Sell	-0.103 (0.044)	-0.105 (0.032)	-0.002 (0.042)	-0.103 (0.031)	-0.105 (0.056)
Sell	-0.118 (0.031)	-0.139 (0.050)	-0.021 (0.044)	-0.118 (0.030)	-0.139 (0.057)	-0.021 (0.057)
Hold	-0.091 (0.013)	0.007 (0.021)	0.098 (0.017)	-0.091 (0.009)	0.007 (0.018)	0.098 (0.019)
Buy	0.011 (0.015)	0.134 (0.019)	0.123 (0.014)	0.011 (0.011)	0.134 (0.010)	0.123 (0.011)
Strong Buy	0.112 (0.018)	0.243 (0.030)	0.131 (0.016)	0.112 (0.012)	0.243 (0.023)	0.131 (0.019)
(Strong Sell)*Affiliation	-0.195 (0.216)	-0.838 (0.320)	-0.642 (0.463)	-0.195 (0.226)	-0.838 (0.266)	-0.642 (0.333)
(Sell)*Affiliation	0.094 (0.136)	-0.087 (0.184)	-0.180 (0.282)	0.094 (0.233)	-0.087 (0.275)	-0.180 (0.339)
(Hold)*Affiliation	-0.001 (0.038)	0.005 (0.050)	0.006 (0.066)	-0.001 (0.040)	0.005 (0.077)	0.006 (0.076)
(Buy)*Affiliation	-0.068 (0.061)	0.013 (0.054)	0.081 (0.088)	-0.068 (0.040)	0.013 (0.029)	0.081 (0.042)
(Strong Buy)*Affiliation	-0.129 (0.031)	-0.023 (0.032)	0.106 (0.049)	-0.129 (0.036)	-0.023 (0.048)	0.106 (0.047)
Sample size	86,962	86,962		86,962	86,962	
R ²	0.0034	0.0085		0.0034	0.0085	

Panel B. Alternative Trade Size Groups

OLS regressions of normalized trade imbalance. Trade reaction is summed over event days 0 and 1. The recommendation levels (Strong Sell, Hold, Buy, Strong Buy) and Affiliation are dummy variables. Sample period is 2/1994 through 7/2001. Standard errors (in parentheses) are robust to heteroskedasticity and arbitrary within recommendation-date correlation across firms.

	Trade Size - Dollar Value (thousands)				
	≤ 5	5-10	10-20	20-50	≥ 50
Strong Sell	0.014 (0.046)	-0.125 (0.044)	-0.112 (0.044)	-0.112 (0.045)	-0.018 (0.045)
Sell	-0.001 (0.040)	-0.095 (0.040)	-0.145 (0.041)	-0.145 (0.037)	-0.012 (0.037)
Hold	0.065 (0.012)	0.006 (0.012)	-0.009 (0.012)	-0.009 (0.011)	0.008 (0.011)
Buy	0.129 (0.011)	0.107 (0.011)	0.070 (0.011)	0.070 (0.011)	0.057 (0.011)
Strong Buy	0.182 (0.012)	0.198 (0.012)	0.145 (0.012)	0.145 (0.012)	0.101 (0.012)
(Strong Sell)*Affiliation	-0.695 (0.288)	-0.465 (0.240)	-0.755 (0.331)	-0.755 (0.256)	-0.199 (0.256)
(Sell)*Affiliation	-0.105 (0.269)	-0.248 (0.204)	0.033 (0.238)	0.033 (0.261)	-0.360 (0.261)
(Hold)*Affiliation	0.088 (0.054)	-0.089 (0.050)	0.003 (0.048)	0.003 (0.046)	0.057 (0.046)
(Buy)*Affiliation	0.013 (0.038)	0.008 (0.037)	0.008 (0.036)	0.008 (0.035)	-0.080 (0.035)
(Strong Buy)*Affiliation	0.059 (0.041)	-0.041 (0.040)	-0.029 (0.037)	-0.029 (0.037)	-0.080 (0.038)
Sample size	86,963	86,963	86,963	86,963	86,963
R ²	0.0073	0.0062	0.0035	0.0035	0.0017

TABLE VI. Analyst Quality

OLS regressions of normalized trade imbalance. Trade reaction is summed over event days 0 and 1. The recommendation levels (Strong Sell, Sell, Hold, Buy, Strong Buy) and Affiliation are dummy variables. The Ever-Affiliated Brokerage sample includes only recommendations made by brokerages who issue at least one affiliated and one unaffiliated recommendation. The All-Star Analysts sample is limited to recommendations made by analysts who were listed in Institutional Investor Magazine's most recent October list of top analysts. Sample period is 2/1994 through 7/2001. Standard errors (in parentheses) are robust to heteroskedasticity and arbitrary within recommendation-date correlation across firms.

	Ever-Affiliated Brokerages			All-Star Analysts		
	Large Trade	Small Trade	Difference S-L	Large Trade	Small Trade	Difference S-L
Strong Sell	-0.129 (0.049)	-0.144 (0.068)	-0.015 (0.084)	-0.153 (0.090)	-0.152 (0.116)	0.001 (0.147)
Sell	-0.153 (0.041)	-0.200 (0.056)	-0.047 (0.070)	-0.260 (0.106)	-0.153 (0.168)	0.107 (0.199)
Hold	-0.098 (0.012)	-0.001 (0.015)	0.097 (0.019)	-0.120 (0.024)	-0.010 (0.029)	0.110 (0.038)
Buy	0.013 (0.012)	0.147 (0.014)	0.134 (0.019)	0.001 (0.027)	0.145 (0.029)	0.144 (0.040)
Strong Buy	0.121 (0.014)	0.270 (0.015)	0.149 (0.020)	0.157 (0.031)	0.353 (0.035)	0.197 (0.047)
(Strong Sell)*Affiliation	-0.170 (0.257)	-0.800 (0.336)	-0.630 (0.423)	-0.517 (0.640)	-1.592 (0.679)	-1.074 (0.933)
(Sell)*Affiliation	0.128 (0.254)	-0.026 (0.274)	-0.155 (0.374)	0.713 (0.505)	0.510 (0.545)	-0.203 (0.743)
(Hold)*Affiliation	0.005 (0.045)	0.013 (0.057)	0.008 (0.072)	0.041 (0.077)	0.116 (0.105)	0.075 (0.130)
(Buy)*Affiliation	-0.070 (0.034)	0.001 (0.039)	0.071 (0.052)	-0.056 (0.059)	-0.051 (0.070)	0.005 (0.092)
(Strong Buy)*Affiliation	-0.138 (0.036)	-0.050 (0.042)	0.088 (0.055)	-0.150 (0.067)	-0.156 (0.086)	-0.006 (0.109)
Sample size	68,948	68,948		11,882	11,882	
R ²	0.004	0.0104		0.006	0.0135	

TABLE VII. Return Analysis

Panel A. Event-time Trading and Post-event Returns

Regressions of market-adjusted abnormal buy-and-hold returns in percent (over the period of trading days indicated in parentheses in the first column) on the dollar value of net buy- minus sell-initiated trades (in \$m). Abnormal returns are calculated using a value-weighted market index. The sample is limited to all firms with at least one year of returns following the recommendation. The recommendation sample period is 2/1994 through 7/2001. Standard errors in parentheses.

	Small		Large		Constant	Sample size	R ²
	Trades (in \$m)	Trades (in \$m)	Trades (in \$m)	Trades (in \$m)			
3 months (2, 64)	-0.1564 (0.1034)	0.0046 (0.0020)	0.1423 (0.0719)	85,577	0.0001		
6 months (2, 128)	-0.4197 (0.1546)	0.0085 (0.0031)	-0.0337 (0.1076)	85,577	0.0002		
12 months (2, 255)	-0.4003 (0.2439)	0.0148 (0.0048)	0.7301 (0.1697)	85,577	0.0001		

Panel B. Analyst Recommendations and Portfolio Returns

Daily portfolio returns in percent. The Long Portfolios are formed by purchasing stocks after strong-buy and buy recommendations; the Short Portfolios are formed by selling after strong-sell and sell recommendations. Both types of portfolios are formed separately for unaffiliated and affiliated analyst recommendations and over holding periods of three, six, and twelve months. A position is held until the holding period expires, the corresponding analyst stops covering the stock, or the analyst issues a new recommendation other than a hold. Hold recommendations extend the current position of the stock by (up to) another holding period. All transactions take place at the end of the day of recommendation. MktMinRf, SMB, HML, and UMD represent the market, size, book-to-market and momentum factors respectively. Turnover is the average daily turnover of the given portfolio. The transaction cost estimates are based on the daily turnover of a portfolio and a transaction cost multiplier based on commissions data and the bid-ask spread. The resulting round-trip cost is 1.669% of traded value. Net returns are the gross abnormal return (the intercept from the regression) less transaction costs. Standard errors in parentheses.

	Returns to Buying the Long Portfolio						Returns to Selling the Short Portfolio					
	3 months		6 months		1 year		3 months		6 months		1 year	
	Unaffiliated	Affiliated	Unaffiliated	Affiliated	Unaffiliated	Affiliated	Unaffiliated	Affiliated	Unaffiliated	Affiliated	Unaffiliated	Affiliated
MktMinRf	0.9469 (0.0109)	1.1698 (0.0381)	0.9515 (0.0107)	1.1716 (0.0295)	0.9494 (0.0107)	1.1587 (0.0236)	-0.8626 (0.0237)	-1.1012 (0.1262)	-0.8858 (0.0181)	-1.0557 (0.1138)	-0.9214 (0.0147)	-1.0926 (0.1107)
SMB	-0.3005 (0.0123)	0.4326 (0.0431)	-0.3002 (0.0121)	0.3712 (0.0333)	-0.3043 (0.0121)	0.2877 (0.0267)	0.3404 (0.0268)	-0.8464 (0.1403)	0.3052 (0.0205)	-0.7711 (0.1283)	0.3030 (0.0166)	-0.7838 (0.1250)
HML	0.1271 (0.0164)	0.1611 (0.0573)	0.1439 (0.0161)	0.1448 (0.0443)	0.1611 (0.0161)	0.2146 (0.0355)	-0.2203 (0.0357)	-0.0716 (0.1939)	-0.2539 (0.0272)	-0.1248 (0.1721)	-0.2292 (0.0221)	-0.1678 (0.1670)
UMD	-0.0304 (0.0079)	-0.1243 (0.0277)	-0.0261 (0.0078)	-0.0809 (0.0214)	-0.0017 (0.0078)	-0.0682 (0.0172)	0.0864 (0.0173)	0.6151 (0.0984)	0.0629 (0.0132)	0.4618 (0.0827)	0.0488 (0.0107)	0.4081 (0.0804)
Constant	0.0117 (0.0063)	0.0027 (0.0218)	0.0094 (0.0061)	-0.0105 (0.0169)	0.0096 (0.0061)	-0.0059 (0.0135)	-0.0151 (0.0136)	-0.0348 (0.0714)	-0.0128 (0.0104)	-0.0588 (0.0650)	-0.0097 (0.0084)	-0.0572 (0.0634)
Observations	1891	1890	1891	1890	1891	1890	1891	1690	1891	1867	1891	1879
R-squared	0.93	0.58	0.93	0.7	0.93	0.77	0.68	0.12	0.79	0.12	0.86	0.13
Gross Return	0.0117	0.0027	0.0094	-0.0105	0.0096	-0.0059	-0.0151	-0.0348	-0.0128	-0.0588	-0.0097	-0.0572
Turnover	0.0193	0.0270	0.0105	0.0141	0.0064	0.0085	0.0262	0.0247	0.0146	0.0162	0.0096	0.0119
Transaction Cost	0.0322	0.0450	0.0175	0.0236	0.0107	0.0142	0.0438	0.0412	0.0244	0.0270	0.0161	0.0198
Net Return	-0.0205	-0.0423	-0.0081	-0.0341	-0.0011	-0.0201	-0.0589	-0.0760	-0.0372	-0.0858	-0.0258	-0.0770

TABLE VIII. Independent Analysts

Regressions of normalized small trades imbalance, sum over event days 0 and 1. Standard errors (in parentheses) are robust to heteroskedasticity and arbitrary within recommendation-date correlation across firms.

	(1)	(2)	(3)
Strong Sell	-0.122 (0.054)	1.066 (0.737)	1.018 (0.840)
Sell	-0.157 (0.049)	-0.204 (0.152)	-0.204 (0.153)
Hold	0.005 (0.014)	0.084 (0.053)	0.122 (0.054)
Buy	0.139 (0.013)	0.196 (0.049)	0.243 (0.053)
Strong Buy	0.246 (0.014)	0.253 (0.047)	0.286 (0.050)
(Strong Sell)*NeverAffiliated	-0.002 (0.139)	-1.035 (0.762)	-0.987 (0.862)
(Sell)*NeverAffiliated	0.125 (0.118)	1.780 (0.152)	1.780 (0.153)
(Hold)*NeverAffiliated	0.032 (0.044)	0.102 (0.108)	0.058 (0.111)
(Buy)*NeverAffiliated	-0.075 (0.045)	-0.394 (0.200)	-0.383 (0.232)
(Strong Buy)*NeverAffiliated	-0.087 (0.049)	-0.168 (0.105)	-0.198 (0.108)
Brokers limited to top 50% in sales ¹		X	
Brokers limited to top 50% in employees ²			X
Sample Size	86,962	4,426	3,959
R ²	0.0085	0.0134	0.0175

Standard errors in parentheses. Standard errors are robust to arbitrary correlation within recommendation-date groups.

¹ Sales represents the dollar value of annual sales for the brokerage firm issuing the recommendation, and is used as a broker size control.

² Employees represents the total number of employees for the brokerage firm issuing the recommendation and is used as a broker size control.

72

The Value of Seeking Financial Advice

Mitchell Marsden · Cathleen D. Zick ·
Robert N. Mayer

Published online: 1 May 2011
© Springer Science+Business Media, LLC 2011

Abstract Retirement planning data gathered from an online survey at a large university in October 2009 are used to examine differences in a variety of retirement planning measures between people who have and have not met with a financial advisor. Problems of self-selection and endogeneity are addressed through the use of propensity scores. The study's major finding is that working with an advisor is related to several important financial planning activities, including goal setting, calculation of retirement needs, retirement account diversification, use of supplemental retirement accounts, accumulation of emergency funds, positive behavioral responses to the recent economic crisis, and retirement confidence. Use of a financial advisor was not related to self-reported retirement savings or short-term growth in retirement account asset values.

Keywords Financial advisor · Retirement planning · Retirement preparedness

Introduction

Over the past 25 years, retirement planning in the United States has become less of an institutional and more of an

individual responsibility. The reasons for this change are well known and include a shift from defined benefit to defined contribution retirement programs, longer life spans, and growing uncertainty about the financial solvency of Social Security, Medicare, and even company pensions (Iams et al. 2009; Moloney and Mistretta 2009; Sun Life Financial 2009; Walker 2008).

At the same time that individuals are shouldering increased financial responsibility, the task of retirement planning has become more difficult. Not only do financial products continue to become more complex, but financial market conditions have become more volatile and daunting. This growing market complexity has led some individuals to select investments based primarily on recent historical returns (Clark-Murphy et al. 2009) which may be very shortsighted. At the same time, the ravaging effects of the recent financial crisis and a “lost decade” for equity investors (Weidner 2009) have caused many people to lose confidence altogether in the stock market and in conventional methods of investment planning.

In light of the increase in individual responsibility for retirement planning, some scholars have argued that automatic savings strategies should be expanded to reduce individuals' growing investment burden (O'Neill 2007). Other scholars have argued that it is simply unrealistic to expect individuals to be able to plan effectively for their retirement (Gross 2005; Rotfeld 2008; Willis 2008–2009). Rather, it takes a paid financial professional to navigate the rough waters of retirement planning. Economist Robert Shiller has proposed that the federal government “subsidize personal financial advice for everyone [by] reimbursing qualified private financial advisors for most or, sometimes, all their fees” (Shiller 2009). Before rushing headlong toward providing every consumer with a financial advisor, however, one should ask whether, and

M. Marsden
Longview Financial Advisors, 220 B Rhett Avenue,
Huntsville, AL 3580, USA
e-mail: mfmarsden@gmail.com

C. D. Zick (✉) · R. N. Mayer
Department of Family and Consumer Studies,
University of Utah, 225 South 1400 East,
Room 228, Salt Lake City, UT 84112-0080, USA
e-mail: zick@fcs.utah.edu

R. N. Mayer
e-mail: mayer@fcs.utah.edu

in what respects, advisors improve the welfare of their clients.

While the entire financial planning profession is built on the premise that professional advice will, over the long run, improve a client's financial well-being, some researchers have suggested that financial professionals often act in their own interest and do little to improve the financial situation of their clients (Bergstresser et al. 2007; Bullard et al. 2007; Bullard and O'Neal 2006; Christian et al. 2008). Others note more generously that while there is evidence of a correlation between greater wealth and the use of professional financial advice, it is difficult to establish any direction of influence or causality (Kramer 2009). It is just as plausible that superior financial well-being leads to working with financial professionals as the reverse. Amid these conflicting opinions and conclusions, the current research addresses the question that still stands: Do financial professionals add to the well-being of their clients and, if so, in what ways? In assessing the empirical evidence that addresses this question, the current study explicitly allows for the endogeneity of seeking advice from a professional and retirement planning outcomes.

Literature Review

Previous research on the link between professional financial advisors and financial planning activities and outcomes can be roughly divided into two areas. One focuses on the factors—including wealth—that determine the use of professional financial advice. The second examines the impact of professional advice on financial planning “outcomes,” including wealth accumulation. Within each type of research, there is wide variation in how financial advice and its correlates are conceived and measured, but virtually all of these studies stop short of dealing with the potential problem of endogeneity, namely, that certain financial characteristics could be the cause and/or consequence of consulting an advisor.

Predicting Use of Financial Advice

The most sustained line of research on the factors that influence use of financial advice has been produced by John Grable and So-Hyn Joo (Grable and Joo 1999, 2001, 2003; Joo and Grable 2001). They conceptualized the choice to seek professional financial help as a form of “help-seeking” behavior. In their model, the decision to seek and use financial assistance was separable from the decision of from whom to seek advice (e.g., professionals vs. nonprofessionals such as family members, friends, and co-workers). The results of both decisions have outcomes that, via feedback, can influence these decisions at a later

time, but the outcomes *per se* are not the central focus of their research. Examining their work as a whole, Grable and Joo found that the correlates of the decision of whether to seek help differed markedly from the correlates of whether one sought professional versus nonprofessional advice. Regarding the latter decision, they found psychological (e.g., risk tolerance, financial stress, attitudes toward retirement planning) and behavioral variables (e.g., financial management practices) were at least as important as demographic and economic ones in predicting the type of advisor used. They observed, in addition, that homeownership and financial satisfaction predicted the type of advice sought, factors that could be viewed as antecedents or consequences of help-seeking. Similarly, they presented confidence in one's financial abilities and confidence in one's financial future as “outcomes” of help-seeking, although these variables could be factors that influence the type of advice chosen. Grable and Joo acknowledged the challenge of separating predictors from outcomes of retirement planning based on professional help and establishing the direction of causality, and they concluded that “research is warranted to explore these potential relationships” (Joo and Grable 2001).

Several subsequent studies, all relying on data from the Survey of Consumer Finances, used measures of wealth to predict use of professional financial advice (Chang 2005; DeVaney et al. 2007; Elmerick et al. 2002; Investment Company Institute 2007). Elmerick et al. (2002), for instance, distinguished between the use of professional financial advice for credit and borrowing only, saving and investing only, and for both types of financial activities. They found that the use of a financial planner for saving and investing only was associated with greater household financial assets, and comprehensive advice was associated with both greater assets and higher net worth. While their study examined financial assets and net worth as two of many potential predictors of financial planner use, they acknowledged the inability to determine the direction of causation between, for example, the use of comprehensive financial advice and higher net worth. They, too, suggested that this is “an interesting question for future research” (p. 230).

Finally, one recent study makes use of a sample of 353 distressed consumers to examine what information sources they used when planning for retirement (Kim and Kim 2010). Professional services (defined to include financial planners, lawyers, lecturers, accountants, and counselors) were used by over half of the individuals in this study. The multivariate analyses did not include any measures of financial attitudes or financial knowledge, and the results revealed only limited support for the role of socioeconomic and demographic factors as they relate to seeking professional financial advice on retirement planning issues.

Determining Possible Impacts of Advisors

Professional financial advice could potentially influence one or more of the many steps in the process of financial planning. These steps include setting goals and specific objectives, creating and implementing a financial plan, and monitoring results (Certified Financial Planner Board of Standards 2009). Research that attempts to measure the possible influence of financial advisors can be organized according to the aspect of the financial process that it examines most closely.

Goal-Related Planning Activities

Setting goals and objectives is the foundation of any financial plan, yet few studies document the extent to which financial advisors contribute to this crucial activity. An exception is a survey by Hartford Financial Services (Thomas 2005). In it, 81% of respondents said, “providing a road map of their financial future” and maintaining client “motivation to do what’s needed to meet retirement goals” were major benefits of working with a financial advisor. In a small-scale survey (Investment Company Institute 2007), over 80% of those who had an ongoing relationship with a financial advisor reported having periodic discussions regarding financial goals. Neither study, however, compared the percentages of people who set financial goals with and without the help of an advisor.

Planning Implementation and Planning Outcomes

A few studies—mostly conducted outside the U.S. and not published in traditional, peer-reviewed outlets—have examined the influence of financial advisors on investment plan implementation and/or investment outcomes. Presumably, plan implementation activities, such as diversification, rebalancing, and portfolio turnover, influence investment outcomes. Nevertheless, the distinction between activities and outcomes is important for the examination of advisor effects inasmuch as plan implementation decisions are far more under the control of advisors (and clients) than wealth accumulation and hence may be more appropriate as yardsticks of advisors performance (Campbell 2006; Fischer and Gerhardt 2007).

With respect to the impact of advisors on the financial planning implementation activities of their clients, previous research has yielded mixed results. Hackethal et al. (2010) examined German clients of an internet brokerage firm and compared customers who traded without an advisor to those who relied on either an independent financial advisor or an advisor affiliated with a branch-based bank. To address the potential problem of self-selection (i.e., the possibility that certain client

characteristics increase both the likelihood of working with an advisor and doing a better job of financial planning), Hackethal et al. used an instrumental variable approach. Rather than measuring use of financial advisors directly, they generated predicted values that were presumed to be unrelated to the outcome measures of interest (e.g., trading, turnover, account diversification, risk-adjusted rate of return). When the instrumental variable was included in a multivariate analysis of these outcomes, the authors found that advisors increased trading, turnover, and diversification (by increasing the share of mutual funds relative to directly owned stocks) but did not improve risk-adjusted returns (Hackethal et al. 2010).

Kramer (2009), studying Dutch investors, found several differences between the plan implementation choices of advised and non-advised investors. Specifically, the portfolios of advised investors had more fixed income securities, a higher number of common equity positions, more international securities, more equity mutual funds relative to total equity allocation, and lower volatilities of returns compared to self-directed investors operating without advisors. Also, advised investors typically executed more trades but had a lower percentage of account turnover. Despite these numerous differences in plan implementation activities, Kramer did not find any difference in the account performance of the two types of investors (Kramer 2009).

The most methodologically compelling study of the impact of advisors was conducted by Gerhardt and Hackethal (2009). Using data on the accounts of 65,000 German bank customers, the authors matched advised and unadvised “nearest neighbors” using a propensity score approach. Each of the roughly 7,000 advised clients was matched with a “non-advised twin” who, based on demographic and account information, was just as likely to have met with a financial advisor but in fact had not. This method approximated an experiment in which subjects are randomly assigned to the treatment group of being advised or the control group of not being advised, thereby addressing the potential problems of self-selection and endogeneity and allowing the authors to conduct a strong test of whether financial advisors influence investment activities and outcomes rather than vice versa.

Using the propensity score approach, Gerhardt and Hackethal observed that advisors improved multiple aspects of plan implementation. Advisors increased enrollment in automatic savings plans, lowered the proportion of equity holdings, increased the use of mutual funds rather than direct equity holdings, increased the number of securities held, raised international equity holdings, promoted equal weighting among investment classes, and lowered trading frequency. In terms of investment results, however, advisors only slightly improved risk-adjusted returns, as measured by Sharpe

ratios (0.90 for advised portfolios vs. 0.88 for non-advised portfolios) (Gerhardt and Hackethal 2009).¹

Several U.S.-based studies of advisor impact have addressed the question of whether advisors help or hurt their clients. A study of life insurance purchases compared people operating without an advisor to those advised by either financial planners or broker/dealers (Finke et al. 2009). The researchers found that the amount of life insurance purchased from a financial planner (but not a broker/dealer) was more likely to be appropriate for a client when compared to insurance bought without the help of a financial advisor.

With respect to investments rather than insurance, the evidence of advisor effects suggests a less sanguine conclusion. Several studies suggest that advisors increase investor costs and thereby lower investor returns. Bullard and O'Neal (2006) examined the operating expense ratios of the S&P 500 index mutual funds purchased by advised and non-advised efforts. They argued that, to the extent that advisors help clients shop for the best value among competing investments, the costs should be lower for funds sold through brokers or financial advisors than those sold directly to investors. They found the opposite: although virtually identical to directly sold index funds, the brokered index funds had significantly higher operating expense ratios without any compensating increase in returns. These findings were consistent with those of Bergstresser et al. (2007). Bullard et al. (2007) found that brokered mutual fund share classes had larger gaps between investor performance and reported fund performance than directly purchased shares, a difference the authors attributed to greater timing errors. Together with finding that share classes with the highest loads (commissions) had the worst investor underperformance, the authors concluded that brokers and advisors sell the products that produce the highest commissions and therefore do not promote the interests of their clients.

Other researchers have questioned the ability (if not the motive) of advisors when it comes to properly assessing the

risk tolerance of their clients (Christian et al. 2008; Roszkowski and Grable 2005). To the extent that riskier investments involve higher advisor compensation, advisors would have an incentive to overestimate client risk tolerance and thereby interfere with the achievement of risk-appropriate investment outcomes.

The tumultuous investment markets during 2008 and 2009 could eventually provide a laboratory for testing the value of professional financial advice. Research on the impact of the stock market's precipitous decline and partial recovery is just beginning to appear. One study (Roszkowski and Davey 2010) found that investor risk perception, but not risk tolerance, increased between January 2007 and June 2009. Another large-scale study, conducted for Wells Fargo, revealed that clients, especially women, contacted their advisor more often after the 2008 stock market downturn than before it. To date, no published study has compared the response of advised and self-directed investors in the post-2008 period. A forthcoming paper, however, shows that people working with a financial planner were "92% more likely to maintain an optimal portfolio composition" during the recent economic downturn (Silverblatt 2010).

In sum, the use of an advisor seems to influence plan implementation activities in ways that potentially benefit the client (e.g., risk diversification through mutual funds), but the evidence is much weaker for any impact of advisors on portfolio returns or investor wealth. Research conducted for Charles Schwab Corporation (2007) concluded that 401(k) participants who received some type of assistance with choosing their investments outperformed participants who did not use some form of guidance, but it also noted that similar superior returns could be achieved by choosing target-date retirement funds or following asset allocation models offered by employer-sponsored investment programs (Charles Schwab Corporation 2007).

Other Impacts

As important as account performance and asset accumulation are to financial planning, perceived financial well-being may be the ultimate measure of the effectiveness of financial advisors. After all, what good is a large retirement nest egg if a person feels financially insecure and worries continually about their retirement security? A series of qualitative studies have suggested that advisees value their financial advisors beyond their impact on portfolio accumulations. In one survey, the greatest reported benefit of working with an advisor was contributing to client peace of mind (Thomas 2005). Ninety-four percent of these clients rated themselves as being very- or somewhat-well prepared for retirement, compared to only sixty-five percent of a 2005 nationally representative sample of U.S. adults who described themselves as being very or somewhat confident

¹ These authors also employed a within-subjects time comparison by examining people who changed from being self-directed investors to being advised. Among the investors who received investment advice as of July 2007, 596 had only begun to work with an advisor during the previous 15 months. For those investors who switched to advised accounts during the period of the study, the authors noted a one-time increase in trading activity as part of an initial portfolio restructuring. These investors also demonstrated a number of presumably salutary changes. Their participation in automated savings accounts increased, their use of speculative investment vehicles decreased, and their portfolios became more diversified through increased use of index funds. Even though the pre-post period was fairly short, the time comparison essentially eliminated the problem of endogeneity, and the findings of the pre-post comparison largely paralleled those in the matched group analysis.

they would have enough money to live comfortably through their retirement years (Helman et al. 2005).

In addition to feeling on track to reaching financial goals, peace of mind may include reduction of the mental burden of financial planning. Accordingly, respondents in a national survey (Investment Company Institute 2007, 2008) cited not having enough time to make their own investment decisions and wanting someone to “make sense of the total financial picture” as reasons for seeking the help of professional financial advisors. Thus, any comprehensive assessment of the impact of financial advisors on clients should also include measures of subjective well-being, such as retirement confidence (Helman et al. 2010).

Anticipated Research Contributions

The study presented here aims to contribute to the existing literature on financial help-seeking behavior in three ways. First, it uses Grable and Joo’s (1999) comprehensive, five-stage framework of personal finance help-seeking behavior but expands on its treatment of “outcomes.” Whereas Grable and Joo (2003) studied several measures of financial confidence, worry, and knowledge as possible outcomes of working with a professional financial advisor, we add to the list of possible outcomes. In particular, this study focuses on multiple stages in the retirement planning process and makes use of self-reported data as well as administrative retirement account data.

Second, the current study accounts for self-selection bias and endogeneity issues by employing a propensity score methodology similar to that employed by Gerhardt and Hackethal (2009). The potential problem of self-selection is severe inasmuch as advised individuals are clearly different from those who are non-advised in ways that could influence their financial planning activities and outcomes independent of any advisor influence. Research has found that individuals who seek professional advisors are, typically, older, married, wealthier, more likely to be female, more likely to be homeowners, and more risk tolerant (Chang 2005; DeVaney et al. 2007; Elmerick et al. 2002; Grable and Joo 1999, 2001; Investment Company Institute 2007; Joo and Grable 2001; Wells Fargo 2009). Endogeneity is an important threat to internal validity as well. The repeated finding that people who consult advisors tended to have higher account balances highlights the possibility that people seek out advisors *after* managing their investments well, receiving an inheritance, or otherwise accumulating financial resources.

Finally, the current study adds to the existing literature by examining the value of a financial advisor in the context of the financial market turmoil of 2008 and 2009. Perhaps the lack of strong evidence in previous research that the accounts of advised individuals outperformed self-directed

accounts is a function of past research being conducted in an era of rising financial markets; it is difficult for advised clients to beat non-advised investors when the value of practically everything is rising. Similarly, it is not difficult for both advised and non-advised investors “to stay the course” during economic good times, but it takes a strong constitution to do so during a market freefall. Thus, it is instructive to examine the potential impact of professional financial advisors during a financial crisis and deep recession.

Methods

Framework

In the context of retirement planning, we used the comprehensive, five-stage model provided by Grable and Joo (1999) to conceptualize the decisions of whether and from whom to seek financial help. Implicit in their model is the argument that individuals assess the benefits and costs of engaging in a range of retirement planning activities, including help-seeking. These benefit-cost calculations are influenced by characteristics of a person’s retirement plan, financial attitudes (e.g., risk tolerance), financial knowledge, and socio-demographic and economic characteristics. Grable and Joo’s model also recognized that financial planning activities have “outcomes,” both objective and subjective. The current analytic approach differed from their model in that (a) it merged the decisions of whether to seek financial advice and from whom into one measure, and (b) it allowed for the possibilities of self-selection and endogeneity associated with estimating the relationship between seeking financial advice and financial planning outcomes.

For individuals covered by an employer-sponsored or employer-administered retirement plan, the characteristics of this plan were hypothesized to affect the likelihood of seeking advice from a professional as well as the engaging in a variety of other retirement planning activities and achieving a variety of retirement-related outcomes. For instance, individuals who are in a defined-benefit retirement plan have less control over and responsibility for their eventual retirement income than people in a defined-contribution retirement plan. As a result, people in defined benefit plans were hypothesized to be less likely to consult with a financial advisor and less likely to engage in retirement planning activities, such as calculating the amount of money that one needs to have saved by the time of retirement.

According to Joo and Grable (2001), an individual’s financial attitudes and psychographic characteristics influence decisions about seeking professional financial advice.

In their model, financial attitudes were measured in terms of risk tolerance, self-esteem, and perceptions of the retirement planning process (e.g., perceptions of the time and effort involved). In our analyses, the focus was on risk tolerance, financial self-efficacy, perceptions of the retirement planning process, and perceptions of the economy and the investment environment. Many of these variables likely reflect the individual's discounting of future benefits relative to planning costs incurred today. As an example, an individual who agrees with the statement, "anyone can have a comfortable retirement if they just plan and save," is likely applying a low rate of discount to the anticipated benefits of planning efforts. Consequently, someone who agrees with the statement may be more likely to meet with a financial advisor and become involved in financial goal-setting and plan implementation than someone who disagrees with the statement.

Financial knowledge also plays an important role, according to Grable and Joo (1999, 2001). On the one hand, those individuals with more knowledge of personal finance, whether it is measured by the answers to objective (correct-incorrect) or subjective (self-assessed) questions, may be less likely to seek professional advice as their knowledge may substitute for such advice. But, at the same time, more knowledgeable individuals may have a relatively greater understanding of benefits conferred by meeting with a retirement advisor. While the relationship between financial knowledge and help-seeking behavior may be complex, it was posited that more financial knowledge is associated with better retirement planning outcomes.

Finally, socio-demographic characteristics such as household income, marital status, age, and education were hypothesized to affect an individual's retirement planning benefit-cost calculation. For example, holding constant plan characteristics, attitudes, and knowledge, younger individuals should be less inclined to meet with a financial advisor and to engage in goal-related and plan implementation activities than older individuals because younger individuals will more heavily discount the benefits of devoting time to such an activity given their relatively longer time frame.

Figure 1 summarizes the framework that was used to guide the research questions. The main focus of the empirical work was to ascertain if those individuals who meet with a financial advisor generally engage in more positive retirement planning activities and achieve more positive financial outcomes than do otherwise similar individuals who do not meet with a professional financial advisor. That is, were the parameter estimates associated with the arrows going from advice seeking behavior to various retirement outcomes positive and statistically significant when accounting for the possible roles of self-selection and endogeneity?

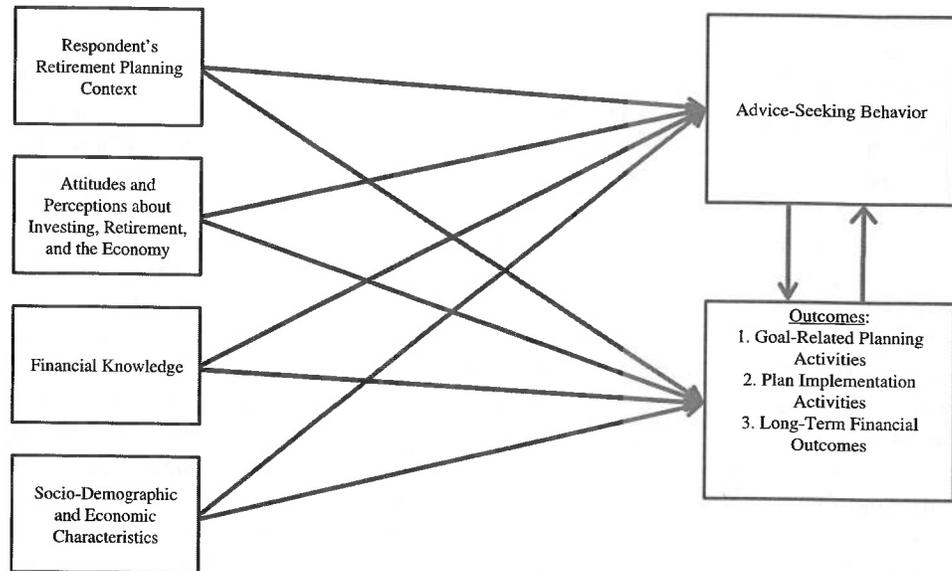
Data Analytic Approach

Concerns about self-selection and possible simultaneity would disappear if individuals in a study could be randomly assigned to either meet with an advisor (i.e., be a member of the "treatment group") or not (i.e., be a member of the "control group"). In practice, individuals self-select into either meeting or not meeting with a financial advisor and that self-selection may be related to their other retirement planning activities such as establishing long-term retirement goals, diversifying retirement investments, and/or placing a relatively high percentage of their investments in the stock market. One way to deal with the possible endogeneity would be to estimate the choice to meet with a financial advisor along with the other retirement planning outcomes as a recursive or simultaneous system. Yet, this approach is limited by the functional form that is chosen and by the reality that such methods may hide the fact that some individuals in the "treated" sample have no counterfactual in the "control" sample (i.e., there is a lack of common support) (Black and Smith 2004; Gibson-Davis and Foster 2006).

Rosenbaum and Rubin (1983, 1984) proposed the use of the propensity score method to address simultaneity problems by balancing members of the treatment group (i.e., individuals who have met with a financial advisor) with specific members of the control group (i.e., individuals who have NOT met with a financial advisor) with regard to their covariates. Essentially, the propensity score adjusts for the bias that is caused by the self-selection of meeting with a financial advisor by creating matches between members of the treatment and control groups rather than through the random assignment that is used in true experiments (Rosenbaum and Rubin 1983, 1984).²

² Three conditions must hold when using the propensity score approach. First, once one controls for observable covariates, potential outcomes must be independent of the treatment selection. This is known as the conditional independence assumption (CIA). In the current application, this means that meeting with a financial advisor should be random once covariates are held constant. In an attempt to meet the CIA, our analyses included variables that captured elements of the four conceptual categories identified in the framework: (1) retirement plan characteristics, (2) attitudes and perceptions regarding retirement planning and the larger financial marketplace, (3) objective and subjective assessments of financial knowledge, and (4) socio-demographic and economic characteristics of the respondent. The second condition that must be met is the common support assumption. That is, the estimated probabilities of meeting with a financial advisor for the treatment group must overlap with the estimated probabilities of meeting with a financial advisor for the control group and the probabilities have to be positive, irrespective of the covariate values (Caliendo and Kopeinig 2008; Imbens 2004; Smith and Todd 2005). In the current study, the trimming method was used to define the common support region (Smith and Todd 2005) which insured that individuals who reported meeting with a financial advisor had a counterpart who did not meet with a financial advisor but who

Fig. 1 The relationship between seeking advice from a financial planner and retirement planning outcomes



The propensity score approach relies on first estimating a logit equation where the dependent variable is whether the respondent has met with a financial advisor at least once in the past two years (1 = yes, 0 = no). The independent variables in this logit model include all factors that might affect the decision to meet with a financial advisor as well as those factors that might affect the substantive outcomes of interest (e.g., setting long-term retirement savings goals, establishing a supplemental retirement account, total retirement savings to date). From the logit estimates, the predicted probabilities of meeting with a financial advisor were generated for all respondents. These predicted probabilities became the features on which treatment respondents were matched to controls.

Once the common support area was identified from the estimated probabilities, members of the treatment group were matched to members of the control group. A number of matching methods have been used in the literature (Gibson-Davis and Foster 2006). Given that there is no consensus on the best matching method, a radius caliper matching technique which makes use of all members of the control group within a 0.05 radius of the treatment observation was used for matching. After the matching was done, *t* tests were conducted to ascertain if statistically

significant differences existed between the treatments and the controls. All propensity score analyses were conducted using PSMATCH2 in Stata 9.0 (Leuven and Sianesi 2003).

Sample

Data for this study were collected in October 2009 through an online survey of all benefits-eligible employees (faculty, staff, and administrators) of a large Mountain West university. Conducted with the cooperation of the university's benefits department and approved by the university's institutional review board, the survey had multiple research objectives of which examining the role of financial advisors in retirement planning was one.

As of September 2009, the university had 15,174 benefits-eligible employees, but only about two-thirds of employees had university-based email addresses. An email invitation to participate in the survey was sent to these 10,152 employees, with 405 being returned as undeliverable. Hence, a maximum of 9,747 employees had the opportunity to read the invitation and decide whether or not to participate. As an incentive to participate, respondents were able to enter a drawing with five iPod Nanos and twenty \$50 restaurant gift certificates as the prizes.

The publicity efforts and participation incentives resulted in a healthy cooperation rate considering the length of the questionnaire and the potentially anxiety-provoking nature of the survey's topics. Recall that this was a year after the stock market's downward plunge in October 2008 and it was a period during which people joked about their 401(k)s having become 201(k)s. By the end of the month (October 2009), 3,131 people submitted completed questionnaires, which yielded a cooperation rate of 32.1%.

Footnote 2 continued

nonetheless had the same estimated probability of having met with an advisor. The final condition that must be met is the stable unit treatment value assumption (SUTVA). SUTVA requires that the outcome for any individual depends on her/his participation only and not on the choice to see or not see an advisor made by any other individual in our sample. Given that our sample was unlikely to have more than one individual from the same household, SUTVA was likely satisfied in the current analyses.

Employer-sponsored retirement plans varied among survey respondents as did supplemental retirement plans. Most faculty members and administrators are enrolled in a defined contribution plan where the university contributes 14.2% of the employee's salary each pay period, but makes no university contribution toward any supplemental plans. This is called the "Standard Benefit Plan" by the university. Participants have a choice between TIAA-CREF and Fidelity Investments as their independent retirement investment providers.

Most staff members are enrolled in a state-run, defined benefit retirement plan (State Retirement System or "SRS"). While this plan provides a benefit based on earnings and years of service, it also contains an additional small defined contribution portion under which the university contributes 1.5% of the individual's salary to a separate, self-directed account. The investment options available to the participant are managed by the State Retirement System; there is no choice among independent retirement providers. Also under this retirement plan, there is no university matching of any supplemental retirement accounts.

A third plan, Hospital Plan Plus (HPP), covers all employees of the university's hospitals and clinics hired after January 1, 2001 and many of the hospital and clinic employees hired before that date. HPP is a defined contribution plan based solely on a university contribution of 6% of salary. There is no employee matching in this defined contribution plan. Beginning approximately a year before the survey, the university began a matching program for supplemental retirement accounts for employees in this third program only.

In terms of the characteristics of the survey respondents, roughly 60% were faculty and administrators (i.e., "exempt" employees) and 40% were "non-exempt" staff (e.g., secretaries, groundskeepers, custodians, nurses). Sixty-five percent of the respondents were women and the median respondent age was 44 years. Approximately 72% of the respondents were enrolled in one of the two defined contribution retirement plans with the remaining 28% participating in the SRS defined benefit plan. As a point of comparison, as of October 2009, 58% of all university employees were women, the median employee age was approximately 42, and 66% of all benefits-eligible employees were enrolled in one of the two defined contribution plans. Thus, the survey respondents generally reflected the larger population of university benefits-eligible employees in terms of gender, age, and retirement plan type.

The primary data source for the current analyses was the online questionnaire. It contains questions covering retirement-related perceptions, attitudes, knowledge, intentions, and behavior, plus socio-demographic characteristics. The

survey included 77 questions, some of which had multiple parts. The mean time spent completing the survey was 22 min.

Retirement account data for most participants in the defined contribution plans managed by TIAA-CREF or Fidelity Investments were linked to the survey respondents who had either employer sponsored retirement accounts and/or supplemental retirement accounts with these vendors. Respondents were informed that their survey data would be linked to information held by the university's benefits office. The data linking was conducted by the university's benefits office to insure data security, confidentiality, and respondent anonymity. In accord with the protocol approved by the university's institutional review board, all identifying information was removed from the data file used by the researchers.

For the analyses presented here, there were two samples. The first sample included those respondents who had no missing data on any of the dependent or independent variables ($N = 2,191$). In moving from the full sample to the sample with no missing data, respondents were most commonly lost for one of two reasons: 315 respondents declined to report the changes they had made in response to the recession and another 494 respondents declined to report their retirement savings.³ This sample is labeled the "Retirement Savings Sample." The second sample is restricted to those respondents who participate in one of the two defined contribution plans offered through the university and who provided information that allowed for linkage to administrative retirement records ($N = 1,319$). This sample was termed the "administrative records sample." In addition, for three of the outcomes that rely on administrative records, this latter sample was further restricted to those 835 respondents who had positive retirement account balances at the end of 2006. This restriction allowed for the calculation of their annual account growth rates over time.

Measures

The variables used in this study fell into three general categories: (1) use of a professional financial advisor, (2) possible impacts of having worked with an advisor ("outcome measures"), and (3) factors that could influence the probability of consulting a financial advisor and/or the possible impacts of consulting an advisor ("independent variables"). Measures of some of these variables were relatively well developed (e.g., retirement confidence),

³ Analyses done with the sample that included the 494 respondents who did not report their retirement savings generated the same results as those done with the sample that excluded them. The results for the former sample are available from the authors upon request.

while others vary substantially from study to study (e.g., use of a professional financial advisor). For the sake of parsimony, the measures used in the current study are discussed as part of the descriptive results.

Results

Descriptive Results

According to Grable and Joo's framework (Grable and Joo 1999; Joo and Grable 2001), help-seeking behavior is influenced by demographic, socioeconomic, psychographic, attitudinal, knowledge, and stressor factors. A correspondingly wide variety of independent variables were used here, and Table 1 shows the descriptive information for these variables for the retirement savings sample and the administrative samples mentioned above. On balance, the descriptive information was quite similar across the two samples.⁴ The modal respondent was female, married, included a spouse and/or another dependent when planning for retirement, participated in the university's standard benefit plan, and owned (or is in the process of buying) a home. S/he was 43 years old, had worked at the university for 10 years, had slightly more than a college education, and lived in a household with an annual income of \$90,290. In only about one-third of the cases did the respondent have a retirement plan from a previous employer and in only a little more than one-third of the cases did the respondent have a spouse who had a retirement plan. Consequently, the majority of respondents anticipated that their employer-administered retirement savings plan would be a, if not *the*, major source of income during retirement.

Table 1 includes four measures of financial knowledge. One is self-rated knowledge used in several studies conducted by the National Foundation for Credit Counseling (NFCC), including its 2009 Financial Literacy Survey. The exact question is: "On a scale from A to F, what grade would you give yourself in terms of your knowledge about personal finance?" The typical respondent rated her/his

financial knowledge as either very good or excellent. These results are quite similar to NFCC's most recent survey (Harris Interactive 2009).

A second question sums the number of different types of information sources that a person consulted in making investments and planning for retirement (AARP 2007). A third question measured whether a respondent knew that his/her employer offered the opportunity to meet free of charge with a financial advisor or counselor about their retirement plans. Respondents were divided into those who answered correctly that advising sessions were available and those who either said sessions were not available or did not know.

A final knowledge measure concerned the fiduciary duties of some but not all types of financial advisors. Four types of advisors were named, and respondents were asked whether each type "pledges to put your financial interests before his or hers." Again, respondents were scored as either answering the question correctly or not. While absent in previous studies of advisor choice and effective, a measure of this type of knowledge would seem to be important in decisions about using advisors, the nature of interactions with them, and the results of using them.

Several types of financial attitudes are potentially relevant to help-seeking behavior and retirement planning outcomes. Joo and Grable (2001) demonstrated that risk tolerance and attitudes toward retirement planning activities predicted the use of professional financial advice. In addition to these two variables, a measure of self-efficacy and perceptions of the future course of the economy were included. The perceptions held by members of our sample as reported in Table 1 were similar to those reported in national surveys. For example, opinions about the ability of planning to ensure a comfortable retirement and perceptions that retirement planning takes too much time and effort were both similar to those expressed in the 2009 EBRI survey from which these items were adapted (Employee Benefit Research Institute 2009). Feelings about the future state of the national economy also matched well with those in a national study (MetLife Mature Market Institute 2009). Respondents in our survey were somewhat less confident, however, in their ability to select investments that best fit their retirement needs than were respondents in a 2007 national poll from which the question was taken (AARP 2007). This difference may have been a function of the fact that this question was asked of our respondents after the economic downturn in 2009 while the AARP survey question was asked before the economic downturn began in late 2007.

Measuring use of professional financial advice was a complex task. Grable and Joo (1999) conceptualized this as a two-step decision, one involving whether to seek help at all and a contingent one to seek advice from a professional

⁴ In moving from the retirement savings sample to the more restrictive administrative records sample, very few of the descriptive statistics change. Most notably, respondents were slightly less likely to be female, educational attainment increased by about 0.7 of a year, and average household income rose to \$101,310. Respondents in the more restrictive sample were moderately more likely to know about the opportunity to meet with a financial advisor free of charge through the university and they were slightly more likely to be future oriented than respondents in the retirement savings sample. They were also more likely to have a retirement plan from a previous employer and to have viewed their employer-administered retirement plan to be a major source of income during retirement. On all other dimensions, the two samples were virtually identical.

Table 1 Descriptive statistics for the variables used in the logistic regression

Variable	Definition	Retirement savings sample (N = 2,233)	Administrative records sample (N = 1,342)
Socio-demographic and economic characteristics			
Gender	1 = female, 0 = male	0.63	0.58
Education	Years of formal schooling	16.7	17.4
Age	Respondent's age in years	43	44
Household income	Annual household income (in \$1,000s)	90.29	101.31
Marital status	1 = married, 0 = otherwise	0.73	0.76
Home ownership	1 = owns or is buying a home, 0 = otherwise	0.82	0.86
Financial knowledge			
Advising knowledge	To the best of your knowledge, does the university offer you the ability to talk free of charge to a financial advisor or counselor about your retirement plans? 1 = yes, 2 = no or don't know	0.78	0.83
Number of sources	Number of different types of resources respondent consults for information regarding investments and retirement planning (range 0–9)	2.5	2.7
Grade	Respondent's assessment of her/his own knowledge of personal finance, 1 = poor, 5 = excellent	3.5	3.6
Fiduciary knowledge	Number of correct answers to questions regarding planner titles (e.g., licensed or registered investment advisor) and fiduciary responsibility (range 0–4)	1.0	1.0
Attitudes and perceptions			
Future orientation	1 = would wait to take a \$1,100 one year from now, 0 = would rather take \$1,000 today	0.43	0.48
Risk	Survey of Consumer Finance 5-point risk tolerance scale, 1 = not willing to take any financial risks, 4 = take substantial financial risks expecting to earn substantial returns	2.4	2.4
Ability	Attitudinal 5-point scale. Statement: I feel confident in my ability to select investments for my retirement plan that are best for my needs (1 = strongly disagree, 5 = strongly agree)	2.2	2.2
Economy	Attitudinal 5-point scale. Statement: I expect the national economic situation to be worse a year from now than it is today (1 = strongly disagree, 5 = strongly agree)	4.1	4.1
Comfort	Attitudinal 5-point scale. Statement: anyone can have a comfortable retirement if they just plan and save (1 = strongly disagree, 5 = strongly agree)	2.31	2.30
Prepare	Attitudinal 5-point scale. Statement: preparing for retirement takes too much time and effort (1 = strongly disagree, 5 = strongly agree)	3.82	3.82
Respondent's retirement planning context			
Years worked	Number of years employed at the university	10.5	10.7
SRS ^a	1 = enrolled in the state defined benefit retirement system, 0 = otherwise	0.25	0.03
HPP ^a	1 = enrolled in the HPP defined contribution retirement system, 0 = otherwise	0.23	0.18
Expected retirement years	Gender-specific life expectancy minus expected age of retirement (measured in years)	13.1	12.9
Others	1 = planning for a spouse or other dependents, 0 = planning for self alone	0.79	0.81
Other plan	1 = respondent has a retirement plan from a previous employer, 0 = otherwise	0.32	0.37
Spouse plan	1 = spouse has a retirement plan, 0 = otherwise	0.37	0.39
Employer retirement major source	1 = employer-administered retirement savings plan will be a major source of income during retirement, 0 = otherwise	0.66	0.72
Supplemental retirement major source	1 = employer-administered supplemental retirement savings plan will be a major source of income during retirement, 0 = otherwise	0.35	0.37

^a The omitted group in this series of dummy variables is composed of individuals who are enrolled in the university's standard defined contribution plan. Individuals can be in the SRS plan and still have TIAA-CREF or Fidelity Investment retirement records if they have a supplemental retirement account with either organization

versus non-professional (e.g., friends, co-workers, relatives) source. In practice, most researchers have collapsed these two steps, as was done here. First, respondents were asked to select their primary source of “information or advice about your retirement investments.” From this list, a professional financial advisor was selected by 25% of respondents.⁵ This percentage was second only to the 34% who selected their spouse or partner as their most frequently consulted source of advice. Then, the 75% of respondents who indicated someone other than a professional advisor as their most frequently consulted source of retirement planning information and advice were asked if they had ever met with a professional financial advisor. This added another 36% for a total of 61% who had ever met with a financial advisor. Then, anyone who had indicated via either question that they had consulted a professional financial advisor was asked to estimate the last time s/he had met with the advisor. Only those who had met with an advisor within the last 2 years were coded as having used a financial advisor. As seen in Table 2, this equated to 42% of the retirement savings sample and 48% of the administrative records sample. This coding decision was based on the assumption that anyone with an active relationship with a financial advisor would have worked with that advisor within the last 2 years.⁶

Regarding goal-related planning activities, the typical respondent in both samples had not established long-term retirement goals nor had s/he estimated her/his retirement financial needs. These findings corresponded to national results (Helman et al. 2009). The likelihood that a respondent had calculated her/his retirement needs increased modestly moving from the retirement savings sample to the administrative records sample (i.e., only those without a defined-benefit plan). The low percentages are noteworthy considering that the dramatic downturn in

⁵ Of these, 43% indicated that they consulted with a professional advisor made available through their employer, while 57% consulted with a professional financial advisor who was *not* made available by their employer. Thus, although all benefits-eligible employees may see a financial advisor free of charge through their employer, many chose to use the services of an advisor who is not affiliated with their employer-provided retirement plan.

⁶ Regarding the job titles of the professional advisors consulted by members of the sample, more than 77% were described as “financial planner or consultant” or “investment advisor.” Other professionals, such as attorneys, accountants, and bankers, represented less than 10% of those named. This refers to anyone who indicated *ever* having met with a financial advisor. In our analyses, the question that asked about meeting with a financial advisor in the past 2 years was used rather than the “ever met” question because the question about meeting with an advisor in the past 2 years is more likely to identify respondents who had an ongoing relationship with a financial advisor. Unfortunately, this meant that the analyses could not distinguish between financial planners, insurance agents, etc., as there was no follow-up question asking about the occupation of the advisor the respondent had seen in the past 2 years.

the economy that occurred in 2007–2008 should have encouraged individuals to take these first, basic retirement planning steps.

There were nine outcomes that refer to plan implementation activities that were measured for both samples. Six of the questions were self-reported behavioral changes “in response to recent economic events.” The vast majority of respondents could not offer an estimate of their basic retirement asset allocation (i.e., stocks vs. bonds vs. cash vs. other). However, roughly a third of the retirement savings sample said they were spending more time learning about financial topics, saving more, and pushed back their expected retirement age in the aftermath of the economic downturn. Seventy-three percent reported using a supplemental retirement account (SRA).⁷ Moving from the retirement savings sample to the administrative records sample, it appears that respondents were more likely to provide an estimate of their asset allocation but were otherwise quite similar to the members of the larger retirement savings sample.

Again, insights about the potential generalizability of the current survey findings were obtained by assessing how the responses in Table 2 compare to the responses from national surveys. Focusing on plan implementation activities, respondents in the current survey generally mirrored respondents in national surveys. For example, 33% of respondents in the survey indicated that they had increased the amount of money that they are saving regularly. In a December 2008 Sun Life Financial survey, 35% of the respondents said they were saving more on a regular basis (Sun Life Financial 2009). Similarly, 16% of respondents in the current survey said that they have increased the amount of money that they owe since the economic downturn while 14% of the respondents in the EBRI survey reported that they had done the same (Employee Benefit Research Institute 2009). Finally, 32% of the respondents in a 2009 national survey (MetLife Mature Market Institute 2009) reported that they had increased the amount of time they spent learning about financial topics while in the current survey the percentage was 30%.

Additional plan implementation information was available for the members of the administrative records sample. While retirement account data such as number of portfolio asset classes or percentage of assets invested internationally were not ideal measures of diversification, they were

⁷ The large percentage with an SRA was likely a reflection of two things. First, recall that individuals in the HPP plan have the option of setting up a 403b plan where the employer will match the employee’s contributions, which likely increased the employee’s incentive to establish and contribute to a SRA. Second, employees at this university may be more financially savvy about retirement matters than the general public given their relatively higher average educational attainment and income.

Table 2 Descriptive statistics for the outcome measures

Variable	Unit of measurement	Retirement savings sample (<i>N</i> = 2,233)	Administrative records sample (<i>N</i> = 1,342)
Advisor last two years	1 = has met with a financial planner within the last two years, 0 = otherwise	0.42	0.48
Goal-related planning outcomes			
Have established long-term goals and working to fulfill them	1 = strongly disagree 2 = somewhat disagree 3 = no opinion 4 = somewhat agree 5 = strongly agree	2.6	2.6
Ever calculated retirement needs	1 = yes 0 = no	0.48	0.53
Plan implementation outcomes			
Could not offer an estimate of asset allocation	1 = yes 0 = no	0.22	0.16
Self-reported use of an SRA	1 = have SRA 0 = no SRA	0.73	0.74
Frequency of reviewing retirement account statements	1 = never or don't know 2 = sometimes 3 = most of the time 4 = every time I receive them	2.2	2.3
Increased the amount of money saved regularly	1 = yes, 0 = otherwise	0.33	0.33
Increased the amount of money owed	1 = yes 0 = otherwise	0.16	0.13
Increased the age of expected retirement	1 = yes 0 = otherwise	0.33	0.32
Increased the investment risk involving existing money in retirement account(s)	1 = yes 0 = otherwise	0.05	0.05
Increased the investment risk involving new contributions to retirement account(s)	1 = yes 0 = otherwise	0.07	0.07
Increased the amount of time spent learning about financial topics	1 = yes 0 = otherwise	0.30	0.30
Percentage invested internationally in 2009	Percentage	na	19
Number of asset classes in 2009	Number	na	2.3
Percentage allocation to stocks in 2009	Percentage	na	64
Percentage of total holdings in small or mid-cap funds in 2009	Percentage	na	10
Percentage in target retirement date funds in 2009	Percentage	na	37
Percentage in actively managed funds in 2009	Percentage	na	79
Long-term financial outcomes			
Self-reported retirement savings	\$	200,793	252,100
Retirement confidence (confidence in having enough money to live comfortably throughout retirement years)	Scale, 1 = not at all confident, 4 = very confident	2.3	2.2
Emergency fund adequacy	0 = no funds 1 = <2 months income 2 = >2 months <1 year's income 3 = >1 year's income	1.6	1.7

Table 2 continued

Variable	Unit of measurement	Retirement savings sample (<i>N</i> = 2,233)	Administrative records sample (<i>N</i> = 1,342)
Total DC account values	\$	na	91,900
Self-reported total supplemental account values	\$	98,206	163,408
Annual rate of account growth over past 3 years including new contributions ^a	Proportion	na	0.18
Annual rate of account growth over past 2 years including new contributions ^a	Proportion	na	0.13
Annual rate of account growth over past 1 year including new contributions ^a	Proportion	na	0.47

^a Based on the 835 respondents who had positive account balances in 2006

suggestive proxies and had the advantage of containing no measurement error because the information was taken from the administrative records provided by TIAA-CREF and Fidelity Investments. As such, they provided tentative insights about diversification. Table 2 reveals that as of the end of the third quarter 2009, the typical respondent had slightly more than two asset classes (out of a maximum of four), 19% of the his/her retirement assets were invested in international stocks, 64% were invested in stocks, 79% held actively managed funds, 10% were in small or mid-cap funds, and 37% were in target retirement date funds.

Finally, turning to the long-term financial outcomes, retirement confidence was gauged using the question asked annually by the Employer Benefits Retirement Institute (Helman et al. 2009). The typical respondent in both samples was somewhat confident that she/he will have enough money to live comfortably throughout his/her retirement years. Respondents in the administrative records sample, on average, had moderately higher levels of self-reported retirement savings and substantially more supplemental savings than those in the retirement savings sample. For the 835 respondents who had administrative records dating back to 2006, the average annual rate of growth in their account over the past 3 years, including new contributions, was 18%. When the time horizon was shortened to the past 2 years, the annual growth including new contributions was 13 percent but over the past year it was 47%. While emergency funds are not a form of retirement saving, the presence of adequate emergency funds was an indication of careful financial planning. The typical respondent in both samples reported having a few months of income in emergency fund accounts.

For long-term financial outcomes, respondents in the current survey typically reported more retirement savings than respondents in national surveys conducted in the past couple of years. In both the 2007 AARP survey and the 2009 EBRI survey, the typical retirement savings for non-

retired individuals were between \$25,000 and \$50,000 (AARP 2007; Employee Benefit Research Institute 2009) whereas the mean retirement savings in our sample was slightly more than \$200,000. This difference likely reflects the relatively generous employer contribution of 14.2% made to employees in the standard defined contribution plan. Although our survey respondents had greater retirement savings on average, they reported levels of retirement confidence similar to respondents in a national survey conducted around the same time (Helman et al. 2009). Overall, a few differences existed between our samples and national samples, but the attitudinal and behavioral data were similar enough to suggest the results of the multivariate analyses that follow are of more general interest.

Multivariate Results

Table 3 contains the parameter estimates of the logistic regressions that were used to generate the propensity scores. Note that separate logits had to be estimated for each sample to insure that the matching was properly executed. Despite differences in the sample restrictions, there were a number of parameter estimates that were remarkably consistent across the equations. Specifically, the likelihood of having met with a financial advisor within the past two years was positively related to age, expected length of retirement, risk tolerance, knowledge of the university's advising opportunities, and the respondent's self-assessed financial knowledge. In addition, respondents in the sample who reported their retirement savings were more likely to have met with an advisor if they were more highly educated, if they thought that the economy will worsen in the coming year, if they believed that anyone can have a comfortable retirement if they plan and save, and if they consulted numerous financial information sources. The parameter estimates of these logits were used to

Table 3 Parameter estimates for the logistic regressions: meeting with a financial advisor within the past 2 years

Variable	Retirement savings sample		Administrative records sample		Administrative records sample with data from 2006 ($N = 835$)	
	Estimated coefficient	Standard error	Estimated coefficient	Standard error	Estimated coefficient	Standard error
Constant	-6.76**	0.70	-6.47*	0.92	-6.52*	1.19
Gender	0.17	0.11	0.13	0.15	0.34 [†]	0.19
Education	0.05 [†]	0.03	0.04	0.03	0.03	0.04
Age	0.04**	0.01	0.05**	0.008	0.06*	0.01
Household income	0.0004	0.001	0.0004	0.002	-0.0001	0.002
Marital status	0.02	0.17	-0.16	0.23	-0.14	0.31
Home ownership	0.29 [†]	0.15	0.11	0.21	-0.24	0.35
Future orientation	0.11	0.10	0.09	0.12	0.005	0.16
Years worked	0.004	0.007	-0.002	0.009	0.002	0.01
SRS	-0.21	0.14	-0.54	0.39	-1.2 [†]	0.63
HPP	0.15	0.13	0.26	0.19	0.43	0.61
Expected retirement years	0.04**	0.009	0.05**	0.01	0.05*	0.15
Risk	0.26**	0.07	0.22*	0.10	0.32**	0.13
Ability	-0.05	0.05	-0.09	0.06	-0.21**	0.08
Economy	0.08**	0.04	0.07	0.05	0.03	0.06
Prepare	-0.07	0.05	-0.10 [†]	0.06	-0.09	0.07
Comfort	0.11**	0.05	0.11 [†]	0.06	0.11	0.07
Advising knowledge	0.71**	0.13	0.92**	0.19	0.93**	0.28
Number of sources	0.06**	0.03	0.04	0.03	0.03	0.04
Grade	0.28**	0.06	0.27**	0.02	0.33**	0.11
Fiduciary knowledge	0.03	0.04	-0.005	0.05	-0.04	0.07
Others	0.13	0.18	0.18	0.23	0.11	0.31
Other plan	0.11	0.11	0.06	0.14	-0.04	0.17
Spouse plan	0.10	0.12	0.18	0.15	0.12	0.18
Employer retirement major source	-0.09	0.05	-0.10	0.14	-0.08	0.20
Supplemental retirement major source	0.14	0.10	0.15	0.13	0.17	0.17
Pseudo R^2	0.13		0.12		0.11	
χ^2	373.77**		215.39**		124.39**	

[†] $p < 0.10$; * $p < 0.05$; ** $p < 0.01$

generate the propensity scores for individuals in both the treatment and the control groups.^{8,9}

⁸ As a check of the CIA assumption, we conducted covariate imbalance testing of the matched means for individuals in treatment and control groups. These t tests, available from the authors upon request, revealed that although we have statistically significant bivariate differences in many of the covariates based on group membership (i.e., having and not having seen a financial advisor), in all cases these differences disappeared once the matching was done. This suggests that the CIA assumption was met in our propensity score analysis.

⁹ Alternative specifications of the logit model were estimated that replaced the income measure with the log of income and the self-assessed financial knowledge measure with a series of dummy variables. This alternative specification allowed for nonlinear effects of income and subjective financial knowledge. There were no

Table 4 shows the means and associated t tests for the outcomes variables for the unmatched and matched samples. Without matching on the probability of having met with an advisor, statistically significant differences were observed between those who had met with an advisor and those who had not with respect to all the goal-related planning activities, about half of the plan implementation and activities, and half of the long-term financial outcomes. Once the probability of meeting with a financial advisor was held constant, some of these relationships disappeared. Specifically, there was no statistically significant difference in self-reported retirement savings, self-reported

Footnote 9 continued
substantive differences in the results when these model modifications were made. The results are available from the authors upon request.

Table 4 Unmatched and average treatment effect on the treated (ATT) using radius caliper matching

Variable		Retirement savings sample			Administrative records sample		
		Treatment	Control	<i>t</i> test	Treatment	Control	<i>t</i> test
Goal-related planning outcomes							
Have established long-term goals and working to fulfill them	Unmatched	2.76	2.42	6.89**	2.84	2.41	6.90**
	ATT	2.76	2.60	2.66**	2.82	2.60	3.08**
Ever calculated retirement needs	Unmatched	0.68	0.34	16.97**	0.71	0.37	13.12**
	ATT	0.68	0.52	6.67**	0.71	0.54	5.50**
Plan implementation outcomes							
Could not offer an estimate of asset allocation	Unmatched	0.12	0.28	-8.83**	0.10	0.22	-5.72**
	ATT	0.13	0.14	-0.74	0.10	0.11	-0.58
Self-reported use of an SRA	Unmatched	0.87	0.63	13.11**	0.85	0.65	8.65**
	ATT	0.87	0.77	5.23**	0.85	0.74	3.86**
Frequency of reviewing retirement account statements	Unmatched	2.34	2.11	5.43**	2.36	2.18	3.51**
	ATT	2.33	2.40	-1.28	2.35	2.42	-1.20
Increased the amount of money saved regularly	Unmatched	0.36	0.31	2.60**	0.37	0.31	2.28*
	ATT	0.36	0.31	2.01*	0.36	0.32	1.33
Increased the amount of money owed	Unmatched	0.13	0.18	-3.00**	0.11	0.15	-2.36**
	ATT	0.13	0.13	0.04	0.11	0.12	-0.18
Increased the age of expected retirement	Unmatched	0.34	0.31	1.44	0.35	0.29	2.34**
	ATT	0.34	0.31	1.23	0.35	0.32	1.18
Increased the investment risk involving existing money in retirement account(s)	Unmatched	0.07	0.03	4.12**	0.07	0.03	2.73**
	ATT	0.07	0.04	2.76**	0.07	0.03	2.49**
Increased the investment risk involving new contributions to retirement account(s)	Unmatched	0.11	0.04	6.07**	0.11	0.04	4.70**
	ATT	0.10	0.05	4.14**	0.10	0.04	3.96**
Increased the amount of time spent learning about financial topics	Unmatched	0.36	0.26	5.36**	0.37	0.24	4.96**
	ATT	0.36	0.27	3.87**	0.36	0.25	3.61**
Percentage invested internationally in 2009	Unmatched	-	-	-	19.03	18.94	0.16
	ATT	-	-	-	19.10	18.30	1.28
Number of asset classes in 2009	Unmatched	-	-	-	2.40	2.26	3.12**
	ATT	-	-	-	2.40	2.28	2.37**
Percentage allocation to stocks in 2009	Unmatched	-	-	-	62.27	65.66	-2.30**
	ATT	-	-	-	62.45	62.23	0.13
Percentage of total holdings in small or mid-cap funds in 2009	Unmatched	-	-	-	10	11	-0.73
	ATT	-	-	-	10	09	0.53
Percentage in target retirement date funds in 2009	Unmatched	-	-	-	31	43	-4.79**
	ATT	-	-	-	31	30	0.41
Percentage in actively managed funds in 2009	Unmatched	-	-	-	77	80	-1.74 [†]
	ATT	-	-	-	77	75	1.77 [†]
Long-term financial outcomes							
Self-reported retirement savings	Unmatched	299,098	127,526	10.85**	340,228	170,945	7.53**
	ATT	297,696	288,084	0.51	335,117	316,472	0.71
Retirement confidence (confidence in having enough money to live comfortably throughout retirement years)	Unmatched	2.90	2.56	9.55**	2.94	2.66	6.47**
	ATT	2.90	2.82	2.07**	2.93	2.85	1.59
Emergency fund adequacy	Unmatched	1.78	1.41	9.47**	1.86	1.49	7.46**
	ATT	1.78	1.65	2.77**	1.85	1.70	2.68**
Total DC account values	Unmatched	-	-	-	-	-	-
	ATT	-	-	-	-	-	-
Self-reported total supplemental account values	Unmatched	154,365	55,599	8.73**	229,164	100,932	7.24**
	ATT	153,050	135,983	1.27	225,458	200,007	1.24

Table 4 continued

Variable		Retirement savings sample			Administrative records sample		
		Treatment	Control	<i>t</i> test	Treatment	Control	<i>t</i> test
Annual rate of account growth over past 3 years ^a	Unmatched	–	–	–	0.17	0.20	–1.41
	ATT	–	–	–	0.17	0.14	0.95
Annual rate of account growth over past 2 years ^a	Unmatched	–	–	–	0.15	0.11	0.80
	ATT	–	–	–	0.15	0.06	1.79 [†]
Annual rate of account growth over past 1 year ^a	Unmatched	–	–	–	0.57	0.38	1.02
	ATT	–	–	–	0.57	0.30	1.32

[†] $p < 0.10$; * $p < 0.05$; ** $p < 0.01$

^a Based on the 835 respondents who had positive account balances in 2006

supplemental account value, the actual percentage of retirement savings held in stocks, the proportion of retirement savings in target retirement date funds, the frequency of reviewing retirement account statements, increasing the amount of money owed, or the respondent's expected age of retirement in response to the economic downturn.

Even after controlling for the possible self-selection of meeting with a financial advisor, a number of statistically significant relationships remained, however. These relationships, although typically somewhat attenuated, suggested that meeting with a financial advisor may encourage individuals to engage in more prudent retirement planning activities. In the case of goal-related planning activities, those who had met with a financial advisor in the past 2 years were more likely to have established long-term goals and to report that they were working to fulfill them. They were also more likely to report that they had calculated their financial needs in retirement. In the category of plan implementation outcomes, those who had met with an advisor were more likely to have a supplemental savings account and to have more asset classes across all of their accounts dedicated to retirement. In terms of responses to the recent economic recession, they were more likely to report that they had increased the amount of money they save regularly, increased the level of risk they take with both new and existing investment funds, and increased the time they spend learning about financial topics compared to otherwise similar survey respondents who had not met with a financial advisor. Finally, turning to the long-term financial outcomes, respondents who had met with a financial advisor reported greater retirement confidence and significantly higher levels of emergency funds.

Discussion and Conclusions

Our multivariate analyses of factors that relate to the likelihood of meeting with a financial advisor are generally

consistent with the predictions derived from Grable and Joo's model (1999). That is, the decision to seek professional advice appears to be a function of financial attitudes and knowledge, and socio-demographic characteristics. Once we control for the potential endogeneity between seeking financial advice and retirement planning outcomes, our research suggests that working with a financial advisor is related to several important financial planning activities, including goal setting, calculation of retirement needs, retirement account diversification, use of supplemental retirement accounts, accumulation of emergency funds, positive behavioral responses to the recent economic recession, and retirement confidence. Use of a financial advisor is not related, however, to self-reported retirement savings or short-term growth in retirement account asset values.

The conclusions drawn from this study should be viewed in the context of the strengths and limitations of the empirical analyses. Perhaps the most important contribution of this study is the allowance that was made for the endogeneity between seeking professional financial advice and retirement planning outcomes. A second strength of the current study is the inclusion of a range of retirement planning activities and outcomes including goal-related activities, plan implementation activities, and long-term financial outcomes. A third strength is the fact that the data were gathered in the aftermath of a deep recession when the counsel of a financial advisor might have greater value. Finally, the use of retirement account data to measure some outcomes enhances the validity of these measures as they are not subject to survey recall error.

Generalizations based on the findings of this study must be made with caution as all survey respondents worked for the same employer. Comparisons of attitudinal responses in this survey to the responses given in national surveys when the same questions were asked suggest that concerns about generalizability should be modest. Yet, differences in the socio-demographic characteristics (e.g., educational

attainment) and retirement plans (i.e., all employees have an employer provided retirement plan and many have an SRA) suggest that one should use caution in extrapolating from this study. In addition, the survey question used to identify whether or not a respondent had met with a financial advisor in the past 2 years did not distinguish among advisor types. Although other information in the survey suggests that the vast majority of respondents who sought professional advice likely met with a professional financial advisor, it was not possible to distinguish in the analyses among the many types of advisors (e.g., stock broker, accountant, financial planner, insurance agent).

In light of the strengths and limitations described above, what has been learned from this study? The empirical work reveals that relatively few of the outcome variables that *might* have been attributed to working with a financial advisor turned out to be so in practice. Some outcome variables were not even related to the use of an advisor in the unmatched sample. Then other differences that existed before correcting for self-selection and endogeneity were screened out in the propensity score analysis. Nevertheless, some outcome variables made it through the various barriers, and these outcomes could be considered among the most fundamental in the retirement planning process.

The foundation of the financial planning process, including retirement planning, is goal setting and the estimation of the amount of money needed to achieve various goals. The use of a financial advisor is positively related to these two outcome variables. The use of an advisor is also related to other actions that are typically regarded as among the first things a person should do in putting their finances on firm ground. One of these is establishing an emergency fund of cash and liquid assets for dealing with unexpected financial events such as a job loss, an illness, or a major automobile repair. Indeed, financial commentator Dave Ramsey calls starting an emergency fund “baby step #1,” even ahead of paying off debt (Ramsey 2010). Similarly, opening an SRA is typically considered the single most savvy action a person can take in terms of retirement planning given the substantial tax advantages of this type of account. Working with an advisor appears to increase the likelihood of taking these two critical steps in this study.

Consulting a financial advisor is related to several other positive financial outcomes. Advisors appear to promote a basic form of diversification, namely, allocating retirement assets to multiple asset classes. Use of an advisor is not related, though, to more subtle forms of diversification, such as buying international stocks or small- and mid-cap stocks. Financial advisors may also help their clients navigate during difficult economic times. In the current study, people who worked with an advisor were more likely to report that they had increased their savings in response to the recession and had stepped up their efforts to learn about

financial topics. Perhaps more important, they also reported having increased the risk level of the money in their existing retirement account as well as new contributions to it. In other words, at a time when many investors were fleeing equities and other high risk investments, advised clients were more likely to take advantage of price reductions in these assets. A final benefit of working with a financial advisor may be increased retirement confidence. This relationship held in the retirement savings sample but not the administrative records sample.

In contrast to the variety of positive financial outcomes that were related to the use of a financial advisor in the current study, several notable outcomes were not. People who consulted a financial advisor reported higher retirement savings in the unmatched comparisons but not after the controls for self-selection and endogeneity represented by the matched samples. Also, among people who provided access to their retirement accounts, there is no difference in the annual rate of account growth between advised and non-advised people when looking at one, two, and three year growth rates. In short, and in line with some previous research (Gerhardt and Hackethal 2009; Hackethal et al. 2010; Kramer 2009), use of an advisor does not measurably increase the amount of money that a person accumulates.

Hackethal et al. (2010) concluded that financial advisors are like babysitters. They provide services to well-to-do people who might be able to do a better job themselves but choose to hire help. Like parents who hire babysitters for their model children, advisors’ clients may mistakenly attribute improved outcomes to their advisor rather than the self selection involved in hiring the financial professional. The results presented here suggest a different analogy. Financial advisors are like clinical psychologists whose services are of value *per se*. They encourage people to examine their most basic desires and priorities, establish concrete goals, withstand (and even taken advantage of) adverse events, and feel confident about their future.

Acknowledgments The research reported in this paper was supported by grants from the University of Utah’s Center on Aging and the University of Utah’s Institute of Public and International Affairs. The University of Utah Benefits Office and Kara Glaubitz also provided valuable assistance with this research project.

References

- AARP. (2007, July). *401(k) Participants’ awareness and understanding of fees*. Retrieved July 25, 2010, from http://assets.aarp.org/rgcenter/econ/401k_fees.pdf.
- Bergstresser, D., Chalmers, J., & Tufano, P. (2007). *Assessing the costs and benefits of brokers in the mutual fund industry*. Boston, MA: Harvard Business School. Unpublished working paper.
- Black, D. A., & Smith, J. A. (2004). How robust is the evidence on the effects of college quality? Evidence from matching. *Journal of Econometrics*, 121(1–2), 99–124.

- Bullard, M., Friesen, G., & Sapp, T. (2007). *Investor timing and fund distribution channels*. University, MS: University of Mississippi, School of Law. Unpublished working paper.
- Bullard, M., & O'Neal, E. S. (2006). *The costs of using a broker to select mutual funds*. University, MS: University of Mississippi, School of Law. Unpublished working paper.
- Caliendo, M., & Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, 22(1), 31–72.
- Campbell, J. Y. (2006). Household finance. *Journal of Finance*, 61(4), 1553–1604.
- Certified Financial Planner Board of Standards. (2009). *What you should know about financial planning*. Retrieved July 25, 2010, from <http://www.cfp.net/Upload/Publications/187.pdf>.
- Chang, M. L. (2005). With a little help from my friends (and my financial planner). *Social Forces*, 83(4), 1469–1497.
- Charles Schwab Corporation. (2007). New Schwab data indicates use of advice and professionally-managed portfolios results in higher rate of return for 401(k) participants. Press Release. Charles Schwab Corporation.
- Christian, J., René, F., & Andreas, H. (2008). *The influence of financial advice on the asset allocation of individual investors*. Wiesbaden: European Business School. Unpublished working paper. European Finance Association 2008 Athens meetings papers.
- Clark-Murphy, M., Gerrans, P., & Speelman, C. (2009). Return chasing as a driver in individual retirement savings investment choices: Evidence from Australia. *Journal of Family and Economic Issues*, 30(1), 4–19.
- DeVaney, S. A., Anong, S. T., & Yang, Y. (2007). Asset ownership by black and white families. *Journal of Financial Counseling and Planning*, 18(1), 33–45.
- Elmerick, S. A., Montalto, C. P., & Fox, J. J. (2002). Use of financial planners by U.S. households. *Financial Services Review*, 11(3), 217–231.
- Employee Benefit Research Institute. (2009). *2009 retirement confidence survey*. Unpublished Survey Results.
- Finke, M. S., Huston, S. J., & Waller, W. (2009). *Do contracts impact comprehensive financial advice?* Lubbock, TX: Texas Tech University. Unpublished working paper.
- Fischer, R., & Gerhardt, R. (2007). *Investment mistakes of individual investors and the impact of financial advice*. Wiesbaden: European Business School. Unpublished working paper.
- Gerhardt, R., & Hackethal, A. (2009). *The influence of financial advisors on household portfolios: A study on private investors switching to financial advice*. Frankfurt: Goethe University. Unpublished working paper.
- Gibson-Davis, C., & Foster, E. M. (2006). A cautionary tale: Using propensity scores to estimate the effect of food stamps on food insecurity. *Social Science Review*, 80, 93–126.
- Grable, J. E., & Joo, S.-H. (1999). Financial help-seeking behavior: Theory and implications. *Journal of Financial Counseling and Planning*, 10(1), 14–25.
- Grable, J. E., & Joo, S.-H. (2001). A further examination of financial help-seeking behavior. *Journal of Financial Counseling and Planning*, 12(1), 55–74.
- Grable, J. E., & Joo, S.-H. (2003). A snapshot view of the help-seeking market. *Journal of Financial Planning*, 16(3), 88–94.
- Gross, K. (2005). Financial literacy education: Panacea, palliative, or something worse? *St. Louis University Public Law Review*, 24(2), 307–312.
- Hackethal, A., Haliassos, M., & Jappelli, T. (2010). *Financial advisors: A case of babysitters?* Frankfurt: Goethe University, unpublished working paper.
- Harris Interactive. (2009, March) *The 2009 consumer financial literacy survey*. Retrieved August 1, 2010, from <http://www.nfcc.org/newsroom/FinancialLiteracy/files/2009FinancialLiteracySurvey.pdf>.
- Helman, R., Copeland, C., & VanDerhei, J. (2009). The 2009 retirement confidence survey: Economy drives confidence to record lows; Many looking to work longer [Electronic Version]. *EBRI.org Issue Brief, April 2009*. Retrieved October 31, 2009, from http://www.ebri.org/pdf/briefspdf/EBRI_IB_4-2009_RCS2.pdf.
- Helman, R., Copeland, C., & VanDerhei, J. (2010). The 2010 retirement confidence survey: Confidence stabilizing, but preparations continue to erode [Electronic Version]. *EBRI.org Issue Brief, March 2010*. Retrieved July 14, 2010, from http://www.ebri.org/pdf/briefspdf/EBRI_IB_03-2010_No340_RCS.pdf.
- Helman, R., Salisbury, D., Paladino, V., & Copeland, C. (2005). Encouraging workers to save: The 2005 retirement confidence survey [Electronic Version]. *EBRI.org Issue Brief, April 2005*. Retrieved May 20, 2010, from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=701302.
- Iams, H., Butrica, B. A., Smith, K. E., & Toder, E. J. (2009). The disappearing defined benefit pension and its potential impact on the retirement incomes of baby boomers. *Social Security Bulletin*, 69(3), 1–28.
- Imbens, G. W. (2004). Nonparametric estimation of average treatment effects under exogeneity: A review. *Review of Economics and Statistics*, 86(1), 4–29.
- Investment Company Institute. (2007). Why do mutual fund investors use professional financial advisers? *Research Fundamentals*, 16(1), 1–8.
- Investment Company Institute. (2008). Ownership of mutual funds through professional financial advisers. *Research Fundamentals*, 17(4), 1–12.
- Joo, S.-H., & Grable, J. E. (2001). Factors associated with seeking and using professional retirement-planning help. *Family and Consumer Sciences Research Journal*, 30(1), 37–63.
- Kim, H., & Kim, J. (2010). Information search for retirement plans among financially distressed consumers. *Journal of Family and Economic Issues*, 31(1), 51–62.
- Kramer, M. (2009). *Investment advice and individual investor portfolio performance*. Groningen: University of Groningen. Unpublished working paper.
- Leuven, E., & Sianesi, B. (2003). PSMATCH2: Stata module to perform full mahalanobis and propensity score matching, common support graphing, and covariate imbalance testing.
- MetLife Mature Market Institute. (2009, March 9). *The 2009 MetLife study of the American dream*. Retrieved July 27, 2010, from http://www.metlife.com/assets/cao/gbms/studies/09010229_09AmDreamStudy_WEB.pdf.
- Moloney, C. X., & Mistretta, C. (2009). *The 2009 Scottrade American retirement study key highlights*. Scottrade (Scottrade o. Document Number).
- O'Neill, B. (2007). Overcoming inertia: Do automated saving and investing strategies work? *Journal of Family and Economic Issues*, 28(2), 321–335.
- Ramsey, D. (2010). *The seven baby steps*. Retrieved August 3, 2010, from <http://www.daveramsey.com/new/baby-steps/>.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for Causal Effects. *Biometrika*, 70(1), 41–55.
- Rosenbaum, P. R., & Rubin, D. B. (1984). Reducing bias in observational studies using subclassification on the propensity score. *Journal of the American Statistical Association*, 79(387), 516–524.
- Roszkowski, M. J., & Davey, G. (2010). Risk perception and risk tolerance changes attributable to the 2008 economic crisis: A subtle but critical difference. *Journal of Financial Service Professionals*, 64(4), 42–53.

- Roszkowski, M. J., & Grable, J. (2005). Estimating risk tolerance: The degree of accuracy and the paramorphic representations of the estimate. *Association for Financial Counseling and Planning Education*, 16(2), 26–47.
- Rotfeld, H. J. (2008). Financial literacy. *Journal of Consumer Affairs*, 42(2), 306–309.
- Shiller, R. J. (2009). How about a stimulus for financial advice? [Electronic Version]. *The New York Times*. Retrieved July 27, 2010, from http://www.nytimes.com/2009/01/18/business/economy/18view.html?_r=1.
- Silverblatt, R. (2010). How to repair your damaged portfolio [Electronic Version]. *U.S. News & World Report Money*. Retrieved May 21, 2010, from <http://money.usnews.com/money/personal-finance/mutual-funds/articles/2010/05/21/how-to-repair-your-damaged-portfolio.html>.
- Smith, J. A., & Todd, P. E. (2005). Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of Econometrics*, 125(1–2), 305–353.
- Sun Life Financial. (2009). *Unretirement index: New results*. Retrieved February 14, 2010, from http://forms.sunlife-usa.com/ga/get_file.cfm?form_id=19983.
- Thomas, T. (2005). Affluent boomers value financial advisors. *National Underwriter Life and Health Financial Services*, 24, 32. July 25.
- Walker, D. M. (2008, February 14). *Addressing fiscal sustainability and fixing the social security System: Two challenges facing the nation*. Retrieved February 14, 2010, from <http://www.gao.gov/cghome/d08524cg.pdf>.
- Weidner, D. (2009). The lost decade of stock investing: Advisers sold us a bill of goods about the lasting value of real estate and stocks [Electronic Version]. *The Wall Street Journal*. Retrieved August 3, 2010, from http://online.wsj.com/article/NA_WSJ_PUB:SB125556534569686215.html.
- Wells Fargo. (2009). *2009 retirement fitness study*. Retrieved August 3, 2010, from http://image.exct.net/lib/ff3016787666/d/1/150007_RDR_Whitepaper%20Brochure_WF.pdf.
- Willis, L. E. (2008–2009). Against financial literacy education. *Iowa Law Review*, 94, 197–285.

Author Biographies

Mitchell Marsden completed his bachelors degree in Consumer and Community Studies at the University of Utah in August 2010. He has taken a position as a financial planner with Longview Financial Advisors in Huntsville, AL.

Cathleen D. Zick is a professor in the Department of Family and Consumer Studies and director of the Masters in Public Policy Program at the University of Utah. She earned her Ph.D. from Cornell University. She is a family economist current working on research in two domains: (1) retirement planning, and (2) the role of time use in promoting healthy body weight.

Robert N. Mayer is a professor in the Department of Family and Consumer Studies who earned his Ph.D. from the University of California, Berkeley. The study of consumer movements, both in the U.S. and elsewhere, is his major research focus. He is currently examining the role of consumer advocates in the creation of the Consumer Financial Protection Bureau.

73



Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Journal of
BANKING &
FINANCE

Journal of Banking & Finance 27 (2003) 1245–1271

www.elsevier.com/locate/econbase

Should you carry the load?

A comprehensive analysis of load and no-load mutual fund out-of-sample performance

Matthew R. Morey *

*Department of Finance and Economics, Lubin School of Business, Pace University, 1 Pace Plaza,
New York, NY 10038, USA*

Received 2 April 2001; accepted 31 October 2001

Abstract

This paper compares the out-of-sample performance of no-load and load mutual funds. Unlike previous studies, this paper provides a more comprehensive analysis as it uses methodologies to incorporate loads directly into the returns. We find two important results. First, after adjusting for loads in the returns data, no-load funds are found to perform much better than load funds, with the differences found to be significant at the 1% level across many different performance metrics. Second, we find that within load funds themselves there is little significant difference in out-of-sample performance between high-load funds and low-load funds even after adjusting for loads.

© 2002 Elsevier Science B.V. All rights reserved.

JEL classification: G23

Keywords: Mutual funds; Loads; Out-of-sample performance; Load-adjusted returns

1. Introduction

In spite of empirical evidence that would seem to suggest that no-load mutual funds outperform or perform at least as well as do load mutual funds, there has recently been a relative increase in the number of load funds that are available to

* Fax: +1-212-346-1796.

E-mail address: mmorey@pace.edu (M.R. Morey).



Fig. 1. Percentage of domestic equity funds that are classified as load funds (1992–2001) where a load fund is defined as a fund that had a deferred or front load.

investors. As shown in Fig. 1, the percentage of domestic equity funds that have a load (front and/or deferred) has trended upward from its low in 1997.¹ This trend towards more load funds is also supported by other developments in the fund industry. For example, the Boston-based Financial Research Corporation stated that in 1995, 45% of the money flowing into mutual funds was invested into no-load, directly sold mutual funds. However, by 2000, this number had dropped to 35% and by 2005 it is predicted to fall to 20%.² Finally, the clearest sign that no-load funds are declining in their importance appeared in 2000 when the 100% no-load Mutual Fund Council (NMFC), the major trade group that represented no-load funds, folded. The NMFC's decline was largely a result of the declining number of diehard no-load funds. Indeed, a new organization has splintered off into a separate group called the Mutual Fund Education Alliance, whose focus is on "investing smartly

¹ Fig. 1 is based on Morningstar's definition of Domestic Equity funds. The graph plots the percentage of Domestic Equity funds that are load funds, i.e., funds that had a front and/or deferred load, for each year starting in 1992 and ending in 2001. The January Morningstar data disks (from 1992–2001) were used to calculate these percentages. Note that the Domestic Equity is the broadest category as it includes diversified US equity funds as well as Hybrid and Specialty funds. Also note that before November 1996, Morningstar did not separate Foreign Equity funds into their own category and instead simply grouped them into the Domestic Equity category. Also, before November 1996, Hybrid funds were given their own category and/or rated differently than other funds.

² The Wall Street Journal, Scudder considers switch away from no-loads, October 16, 2000, page C1.

and with an eye toward costs” rather than the NMFC’s pledge of investing directly without sales charges.³

Why is this increase in load funds occurring? Indeed, many fund experts predicted that with the increase in the popularity of mutual funds exactly the opposite would have taken place. Investors would realize that they could avoid costly load charges by buying no-load funds. Moreover, the Internet would provide many cheap sources of information that would allow investors to monitor funds on their own, without the help of an advisor or broker.

The answer to this question is not simple as there are many issues affecting the industry. However, there are some basic explanations that have been discussed in the press. One of these explanations is that many firms have found it difficult and costly to sell no-load mutual funds directly to the public as the no-load business has become extremely competitive with large firms, like Fidelity and Vanguard, having economies of scale advantages. Indeed, the process of advertising directly to the public is very expensive, ranging from \$300 to \$500 just to lure a new customer to call.⁴ Another explanation for the trend is that there is some anecdotal evidence that customers, now more than ever, want the advice and extra services that come with load funds. Industry experts point out that because the number of funds available to investors has jumped 10-fold in the last decade, and because there is so much more information on funds in general, many investors are asking (and hence willing to pay) for advice and services.⁵ Indeed, Kihn (1996) argues that the majority of mutual fund investors are more concerned with customer services than with financial performance.

While there may be other explanations for this trend towards more load funds, the object of this paper is not to examine why this change is occurring but rather to assess the costs for investors buying load funds as compared to no-load funds. In other words, we ask, what are the costs to investors in terms of fund performance for the advice that usually comes with buying a load fund?

This question already has received some treatment in the literature. Various papers by Carlson (1970), Ippolito and Richard (1989), Elton et al. (1993), Grinblatt and Titman (1994), Droms and Walker (1994), Gruber (1996), and Fortin and Michelson (1995) have examined the performance of load and no-load funds and, in general, they find that before the effect of loads is incorporated there is no significant difference in returns between load and no-load funds.⁶ However, our paper differs

³ Jaffe, Charles, “Load, No-Load a Dead Issue: Consider Costs”, *The Boston Globe*, May 24, 2000, page D4.

⁴ *New York Times*, November 5, 2000, page C8.

⁵ Various recent newspaper articles make this point. See for example, “Death of the Fund Salesman Has Been Greatly Exaggerated”, *The New York Times*, October 8, 2000, “Investments-No Load? No Way! Ten years ago, no-load funds were poised to take over the industry. Today, fund companies proudly add load products”, *Financial Planning*, May 1, 2000, “Scudder Considers Switch Away From No-Loads”, *The Wall Street Journal*, October 16, 2000, and “Load versus No-Load: It’s a question of paying for advice”, *USA Today*, September 1, 2000.

⁶ One paper that comes up with a different result is Ippolito and Richard (1989) who suggests that “Load funds earn rates of returns that plausibly off-set the load charge” (see Ippolito and Richard, 1989, pp. 14–15).

from these others in that it provides the most comprehensive study of this issue to date, incorporating many features that the other papers do not utilize, or if they do, they do not do so in as comprehensive a manner as this paper. Specifically, these features include:

1. A direct incorporation of loads into the out-of-sample returns. As a result we are able to examine the *load-adjusted* performances of load funds and no-load funds. This approach allows us to compare the two types of funds on a level playing field. This method differs from all of the previous papers that have simply examined non-load-adjusted returns and hence could not directly quantify the difference in realized performance between investors in load and no-load funds.
2. Mutual fund data generated at the time the funds were actually available to investors. We then follow the out-of-sample performance of *all* of these funds. This methodology allows us to circumvent the well-known survivorship bias problem that is described by Brown et al. (1992), Elton et al. (1996b) and others. Other than Ippolito and Richard (1989), Elton et al. (1993) and Gruber (1996) our paper is the only load/no-load paper of those listed above that does not have a survivorship bias problem.
3. A large sample of funds that allows for a more precise characterization of performance. Indeed, we examine all of the domestic equity funds that were rated by Morningstar on December 31, 1992. This amounts to 635 funds of which 304 were no-load funds and 331 were load funds. As a result, with the exception of the sample used by Fortin and Michelson (1995), our sample is more than twice the size of any sample used in the other studies. Furthermore, while the Fortin and Michelson sample is larger than ours, it suffers from a severe survivorship bias problem as they sample funds from the January 1994 Morningstar data disk and simply look backwards to 1976 to examine the returns. Any funds that dropped out of the sample from 1976 to 1994 were excluded from their sample.
4. A sample of funds that allows us to examine the load versus no-load performance of funds across different styles and ages of funds. The only other papers to examine style effects are within the load and no-load context is (Fortin and Michelson, 1995). To our knowledge, none of the existing papers has examined age effects in the context of the load/no-load funds.
5. A sample of funds that also allows us to assess whether the size of the load contributes to differences in performance by investors in no-load and load funds. Furthermore, *within* the load fund group we are able to assess whether performance varies from low-load and high-load funds. None of the other papers have examined these types of effects.
6. Four different measures of performance: Mean monthly returns adjusted for the style of the fund, a Sharpe ratio, a Jensen-type alpha and a 4-index alpha. These measures are similar to those used by Gruber (1996) and are slightly different than the two measures (single index and a 3-index model) used in Elton et al. (1993) and the two measures (the 10-factor portfolios and eight characteristic-based

portfolios) used in Grinblatt and Titman (1994). The other related papers generally only use one or two performance measures with usually those being the Jensen's alpha, Treynor index or mean returns.

7. A data set with out-of-sample returns covering the period of 1993–1997. Although not absolutely current, these data are much closer to the present than the data used in the other existing studies.

The rest of the paper is organized as follows. Section 2 extensively describes the data that we use in the paper. Section 3 describes the methodology of the paper, Section 4 presents the results and Section 5 provides the conclusions.

2. Data

To better organize the description of the data, this section is divided into five subsections: Fund selection criteria, the out-of-sample evaluation period, the types of data used, merger and problem funds, and the load-adjusted returns.

2.1. Fund selection criteria

To select funds we use the January 1993 Morningstar On-Disk. This provides data for all funds that were available to US investors as of December 31, 1992. Our rationale in using the 1993 disk rather than a later disk is that it enables us to examine the out-of-sample performance of a smaller number of funds (635) as opposed to well over 2000 funds if we were to use the 1996 or 1997 On-Disk/Principia Data Disks. This lower number of funds is important because the process of tracking each fund's out-of-sample returns through all mergers, name changes and liquidations is quite onerous.

From this disk we then select all open ⁷ funds with at least three years of return history ⁸ that are within each of the following five Morningstar "Investment objectives" (styles): aggressive growth, equity-income, growth, growth and income, and small company. This produced the sample of 635 funds. Our rationale for selecting only open funds is that we wanted all the funds selected to be actually available to investors as of December 31, 1992. We use the three-year history criterion since it ensures that each fund will contain enough in-sample data to calculate some of our performance metrics (see Section 3.3 for more on this issue). Finally, the use of different styles allows us to examine whether or not there is a "style effect" in the out-of-sample performance. In regards to this style effect, it should also be noted here that the five styles used in our analysis are assigned by Morningstar, meaning

⁷ There were 24 funds that meet our other criteria and yet were listed as closed funds in January 1993. These funds were excluded from the sample.

⁸ That is, the inception dates of the funds must begin before December 31, 1989.

that the style designations are much more accurate and up-to-date than the style information provided by the funds themselves.⁹ Moreover, these five styles make up the vast majority of domestic equity funds and the use of these five styles is in the tradition of Blume (1998).

2.2. *The out-of-sample evaluation period*

For each of the 635 funds we then examine the out-of-sample performance for five years, from 1993 to 1997. We use five years since this a common middle-term interval of time for evaluating performance.

2.3. *Types of data used*

With this sample of 635 funds we then acquire the following data for each fund:

1. *The in-sample monthly return history from 1990 to 1992.* This again is available from the January 1993 Morningstar disk. These return data account for management, administrative, and 12b-1 fees and other expenses automatically taken out of fund assets, however they do *not* account for loads.¹⁰
2. *The out-of-sample monthly return history from 1993 to 1997.* This information is taken from later Morningstar data disks (quarterly data disks ranging from 1993 to 1998). As with the monthly returns from 1990 to 1992, these returns account for management, administrative and 12b-1 fees and other costs, but they do *not* account for loads.
3. *The front-load of the fund as of December 31, 1992.* This information is taken from the January 1993 Morningstar data disk.
4. *The deferred load of the fund (as of December 31, 1992) assuming that the fund was held for a period of five years.* This information is taken from the January 1993 Morningstar disk. We only include these deferred loads as we assume that each fund will be held for this length of time. Any deferred load that is applied to holding periods less than five years is not considered (we explain these data more in Section 2.5).
5. *The age and style of the fund as of December 31, 1992.* This information is taken from the January 1993 Morningstar disk.

⁹ We also examined if the funds retained their style classifications by Morningstar in the out-of-sample period. We found that, in our sample, more than 87% of the funds retain their style classification at the end of the sample period. Hence, according to Morningstar, the vast majority of funds did not change their style of management. We did, however, conduct an analysis in which we use the styles of the funds at the end of the out-of-sample period (December 1997). These results were very similar to those using the styles at the beginning of the out-of-sample period. These results are available in Table 5.

¹⁰ See the Morningstar Principia Manual (1998, p. 107). Note also that a number of papers (e.g. Detzler, 1999) that have found that fund expenses are negatively related to performance.

2.4. Problem funds

As described in the previous section, we select funds at the time the funds were listed by Morningstar. To examine the out-of-sample performance, we then obtain the out-of-sample monthly returns of these funds. For a majority of the funds, obtaining the out-of-sample returns is simply a matter of following the fund's future performance. However, because a minority of funds have either gone through a name change, a merger, a combination of both, or because they have liquidated, identifying out-of-sample returns for those funds is more complicated. In this section, we describe how we handle these problematic funds.

For name changes, we use the Morningstar data¹¹ and *The Wall Street Journal* to identify the name changes. We then simply use the renamed fund's returns as the out-of-sample returns.

For the merger funds we used the Morningstar data and *The Wall Street Journal* to ascertain the month of the fund merger. If these two sources did not provide the necessary information, we called the individual mutual fund companies. Once the merger month was identified, we then collected the out-of-sample returns by the following procedure. First, until the fund merged, we simply use the out-of-sample returns of the fund in question. After a fund has merged into its partner fund, we assume the investor randomly re-invests into one of the other surviving funds of the same style *and the same load preference*, i.e., load or no-load.¹² Hence the out-of-sample returns from the merger month and onwards are equally weighted averages of the returns of all the other surviving funds in our sample with the same style *and load preference*, with the load and style status determined as of December 31, 1992. For example, the returns from the merger month onward of no-load growth fund would be the equally weighted average returns of all the other surviving *no-load* growth funds. If instead the merger fund was a growth fund with a *load*, its returns from the merger month onward would be the equally weighted average returns of the surviving load growth funds.¹³

For the liquidated funds we first identify when the fund was liquidated. Again, this information was obtained from Morningstar or *The Wall Street Journal*. As with the merger funds, from the month of liquidation and onwards, we assume the investor randomly re-invests into those funds in our sample with the same investment objective *and load preference* as the liquidated fund.

¹¹ The Morningstar On-Disk and Principia disks (after 1993) both provide a list of funds that have recently undergone name changes, mergers and liquidations.

¹² We classify any fund that has a front load and/or a deferred load (for the five-year holding period) as a load fund. Any other fund is classified as a no-load fund.

¹³ Of course, an alternative approach would be to use the "follow-the-money" approach introduced in (Elton et al., 1996b), where a merged fund's returns are spliced to its "merge partner" fund's returns to form a complete time series. However, we did not use this method since we require a complete in-sample time series of returns, i.e., 1990–1992, for the merge partner fund, and in some cases the partner fund did not exist long enough to obtain such a series.

It should be noted here that as a robustness check, we calculated our results using two additional methods for the merger/liquidation funds. First, we assumed that after a fund merged or liquidated, its returns were an equally weighted average of the returns of the surviving funds of just the same style. For example, the returns from the merger month onward of a growth fund would be the equally weighted average returns of all the other surviving growth funds, regardless of whether they were load or no-load funds. For the second method, we assumed that once a fund merged or liquidated, its returns were an equally weighted average of the returns of the surviving funds that were of the same style *and possessed a 3-star or better overall Morningstar rating at the time the fund merged*. This rule required that we check the Morningstar ratings of the surviving funds at the time of the merger. The reason we incorporated this second additional method was to incorporate a rule that allowed the investor to re-invest their money into a high performing fund of the same style. This being said, both additional merger/liquidation methods were found to produce similar results to those presented in the Section 4. Hence, we do not report these results, however, they are available upon request.

2.5. Returns data and load adjustments

For the out-of-sample returns and the in-sample returns, the data consist of monthly returns from the Morningstar On-Disk and Principia programs. As stated earlier, these returns data are adjusted to account for management, administrative, and 12b-1 fees and other costs automatically taken out of fund assets, however, these monthly return data do not adjust for sales charges such as front and deferred loads.¹⁴ Consequently, if we use the monthly return data for the out-of-sample returns, investor returns on load funds are overstated. The question is, thus, how to incorporate loads into the monthly return data?

Very little attention in the mutual fund performance literature is given to the treatment of loads in return data. Although some authors (e.g., Elton et al., 1993; Grinblatt and Titman, 1994; Gruber, 1996) have presented results separately for load and no-load funds, all the no-load/load context studies provide no direct adjustment for loads in their returns data. As a result, previous studies have only been able to hint at the true difference in returns between load and no-load funds.

Rather than follow previous approaches, we wanted to directly adjust our results for loads by using load-adjusted returns. To do this we use an approach similar to that in Rea and Reid (1998). For front loads, we consider an investor who buys and holds the load shares for our holding period of 60 months (five years). As with most front loads we assume that the investor buying the fund pays a load in a lump sum at the time the fund is purchased. To spread the front load across the period that the shares are held, we use Rea and Reid's assumption that the investor borrows the amount necessary to pay the load up front and then repays the loan as an annuity in equal, monthly installments during the holding period. Hence, the

¹⁴ Morningstar Principia Manual (1998, p. 107).

monthly load adjustment reflects the amount that was borrowed and the interest on the loan.

Mathematically, our front-load adjustment process is the following:

$$f^m = \frac{f}{\sum_{j=1}^h (1+r)^{-j}} \quad (1)$$

where r is the monthly interest rate (the monthly periodic interest rate of five-year Treasury yield in January 1993), f the front-load (expressed as a percent), h the number of months the fund is held and f^m the monthly front-load adjustment.

Hence, the front load-adjusted (for front loads) returns are

$$R_{it}^{\text{FLA}} = R_{it} - f^m,$$

where R_{it} is the non-load-adjusted monthly return of fund i in month t , where t goes from 1 to 60. R_{it}^{FLA} is the monthly front load-adjusted return of fund i in month t .

As an example of the above adjustment, consider a five-year investment in Fidelity's Magellan fund starting in January 1993. As of January 1993, that fund had a front-load of 3%, and the five-year Treasury yield was 5.83%, giving a monthly periodic rate of 0.4853%.¹⁵ Therefore, for the five-year holding (out-of-sample) period, $f = 3\%$, $r = 0.004853$, and $h = 60$, giving $f^m = 0.0577\%$. We then subtract 0.0577% from each of the Magellan fund's 60 monthly returns from 1993 to 1997 to obtain the front load-adjusted returns.

For the deferred load adjustment, the process is different in a number of ways. First, we only define a fund as having a deferred load if the fund has "five-year holding period deferred load" as of December 31, 1992. That is, as of December 31, 1992, the fund must show evidence that even if the investor holds the fund for a period of five years (1993–1997), a deferred load is still charged. Any deferred load that is applied to holding periods less than five years is not considered.

Second, we assume in our analysis that the five-year holding period deferred load that is stated in December 31, 1992 is in fact applied five years later. Any changes in the deferred loads from December 31, 1992 to December 31, 1997 are not accounted for. While at first blush this may seem to be problematic, this assumption is consistent with the general industry practice of "locking-in" the deferred loads at the time the fund is purchased. Of course, funds could have changed their deferred loads during the out-of-sample period, but we found no evidence of such after calling all of the existing deferred load funds in our sample.¹⁶

Third, deferred loads are often calculated differently than front loads. Specifically, deferred loads are sometimes calculated on the value of the fund at the time

¹⁵ These data are from the Federal Reserve Database at www.econmagic.com.

¹⁶ We called 20 mutual funds (31 deferred load funds minus 11 that had merged or liquidated since August 2001). All 20 funds reported that they lock in the deferred load structure at the time the fund was purchased. Moreover, this policy was followed by all 20 funds since December 31, 1992. It should also be noted here that while all 20 funds followed the lock-in policy on deferred loads, the funds could have legally not followed this policy. See National Association of Security Dealers (NASD) rule 2830 and the Investment Company Act of 1940 (Rule 6c-10) for more information.

the investor sells rather than the up-front costs. If the investor has invested very successfully this can mean that the amount paid for a small deferred load can actually be greater than that of a large front load.¹⁷ In our paper we assume that all deferred loads are calculated using the up-front costs because a strong majority of the fund companies that we called also followed this method.¹⁸

Fourth, the last difference between deferred and front loads of course lies in the fact that the payment of the deferred load does not occur until the end of the holding period. To convert the deferred load into a monthly payment, the investor is assumed to prepay the deferred load in equal monthly installments. The amount of the monthly prepayment reflects the deferred load less the interest earned on the prepayments.

Thus the equation for the monthly deferred load adjustment is

$$d^m = \frac{d}{\sum_{j=1}^h (1+r)^j} \quad (2)$$

where d is the deferred load (expressed as a percent), d^m the monthly deferred load (assuming a five year holding period) adjustment.

Hence, the deferred load-adjusted returns are

$$R_{it}^{\text{DLA}} = R_{it} - d^m,$$

where R_{it} is the monthly return of fund i in month t , R_{it}^{DLA} the monthly deferred load-adjusted return of fund i in month t .

Note finally that if a fund has no deferred load for the five-year holding period and no front load, then it is considered a no-load fund.

3. Methodology

To measure out-of-sample performance we use four performance metrics: The mean monthly excess returns, the Sharpe (1966) ratio, a modified version of Jensen's (1968) alpha, and a 4-index alpha. For each performance metric we examine both *non-load-adjusted* and *load-adjusted* versions. We now explain, in detail, the four out-of-sample performance metrics.

3.1. Excess mean monthly returns

The non-load-adjusted excess monthly returns for the i th mutual fund during the out-of-sample period are signified by $R_{it} - R_{ft}$, where R_{ft} , is the 30-day T-bill rate.

¹⁷ We thank an anonymous referee for this point.

¹⁸ Nineteen of the twenty funds that we called reported that they charged the deferred load on the up-front costs. The only fund in our sample that did not use up-front costs was the Quantitative Growth and Income Fund which charged the deferred load on the ending value. It should be noted that this information, about how the deferred load is specifically calculated, is not available in the Morningstar data.

The non-load-adjusted *mean* monthly excess return for the i th mutual fund during the out-of-sample period is $\overline{R_i - R_f}$.

The load-adjusted excess monthly returns for the i th mutual fund during the out-of-sample period are signified by $R_{it}^{LA} - R_{ft}$, where $R_i^{LA} = R_{it} - f^m - d^m$. The load-adjusted *mean* monthly excess returns are simply equal to

$$\overline{R_i^{LA} - R_f}.$$

3.2. The Sharpe ratio

The non-load-adjusted Sharpe ratio is

$$\text{Sharpe}_i = \frac{\overline{R_i - R_f}}{\sigma_i} \quad (3)$$

where σ_i is the standard deviation of $R_{it} - R_{ft}$.

The load-adjusted Sharpe ratio for fund i is

$$\text{Sharpe}_i = \frac{\overline{R_i^{LA} - R_f}}{\sigma_i^{LA}} \quad (3a)$$

where σ_i^{LA} is the standard deviation of $R_{it}^{LA} - R_{ft}$.

3.3. Modified Jensen and 4-index alphas

For two additional alternative predictors, we use Jensen single-index and 4-index alphas. The following time-series regression model is used:

$$R_{it} - R_{ft} = \alpha_i + \sum_{k=1}^K \beta_{ik} I_{kt} + \varepsilon_{it} \quad (4)$$

where $R_{it} - R_{ft}$ is the excess total return (net of the 30-day T-bill return) for fund i in in-sample month t , α_i is the alpha for fund i , used as a performance predictor, β_{ik} is the sensitivity of fund i 's excess return to index k , I_{kt} is the return for index k in in-sample month t and ε_{it} is the random error for fund i in in-sample month t .

For Jensen alphas, $K = 1$ and I_{1t} is the excess total return of the S&P 500 in month t . For the 4-index alphas, $K = 4$, I_{1t} is the excess total return of the S&P 500 in month t , I_{2t} is the excess total return of Lehman Aggregate Bond Index in month t , I_{3t} is the difference in return between a small-cap and large-cap stock portfolio based on Prudential Bache indexes in month t , and I_{4t} is the difference in return between a growth and value stock portfolio based on Prudential Bache indexes in month t .¹⁹ We utilize the 4-index model because, as shown in Elton et al. (1996a), this model provides for better risk adjustment for mutual funds than does the single-index model.

The non-load-adjusted modified Jensen and 4-index alphas are calculated using a methodology similar to that of Elton et al. (1996a). Specifically, we utilize a time

¹⁹ See Elton et al. (1996a) for a detailed description of the Prudential Bache portfolios used in the 4-index model.

series period of monthly non-load-adjusted returns going back three years from the selection date and forward to the end of the out-of-sample evaluation period to obtain an estimate of the intercept from either the single-index or 4-index model regression (Eq. (4)). As mentioned in Section 2, to be included in the sample each fund had to have three years of in-sample returns.

To obtain the alphas, we add the average monthly residual during the evaluation period to the intercept. For example, to obtain the modified Jensen alpha, we run the 1-index model on monthly returns starting in January 1990 and ending in December 1997 (8 years) to obtain an estimate of the intercept. We then add the average of the fund's residuals during the five years after the selection date (1993–1997) to the estimated intercept to obtain the fund's modified Jensen alpha.

To obtain alphas for funds that merged or liquidated during the evaluation period, we proceed as follows. First, we run two regressions: (1) A regression using the fund's returns starting in January 1990 and ending in the month prior to the fund's disappearance and (2) a regression run over the entire sample period (1990–1997) using the returns of an equally weighted portfolio formed each month from the existing funds of the *same style and load preference*.²⁰ We then form a weighted average of (1) the fund's estimated intercept plus the fund's average residual during the time it survived in the evaluation period and (2) the estimated intercept plus the average residual during the remaining time in the evaluation period of the equally weighted portfolio, where the fund's weight is the fraction of the evaluation period it survived and the equally weighted portfolio's weight is the remaining fraction. This provides a performance measure for an investor who buys a remaining fund in the sample at random if the original fund merges or liquidates.

For the load-adjusted modified Jensen and 4-index alphas we use the same methodology described above however, we use the excess load-adjusted returns, $R_{it}^{LA} - R_{ft}$, for the out-of-sample returns. That is, we use the excess non-load-adjusted returns for the in-sample data (1990–1992) and the excess load-adjusted returns for the out-of-sample period (1993–1997). Our rationale for not using load-adjusted returns during the in-sample period is that we assume the investor has not yet bought the fund and hence a load should not be subtracted from the returns. Moreover, the loads may be quite different during the in-sample period than during the out-of-sample period and hence it would be difficult to know what load to apply and for how long to apply it.

4. Results

The results are organized into six subsections: Summary statistics on the sample; overall results; style effects; age effects; size of the load effects and size of the load effects in load funds.

²⁰ Hence for the second regression there are 10 different equally weighted portfolios (five styles and two load preferences). We use the equally weighted portfolio that matches the style of the merger/liquidated fund.

4.1. Summary statistics on the sample

The summary statistics of the sample are presented in Tables 1 and 2. Table 1 shows the number of load and no-load funds, the average front load, the average deferred load, the number of liquidated funds, and the number of merger funds. Table 1 also shows the same information organized across styles and ages of the funds. Table 2 shows the same information organized across the size of the loads. Again, the load numbers, styles and ages are all as of December 31, 1992.

The two tables show several interesting findings. First, 97 of the 635 funds in our sample (about 15%) merged or liquidated during the out-of-sample period, hence survivorship bias is obviously an important issue to consider in this paper.²¹ Second, the average front load was 5.15% with equity-income funds having the highest average front load. The highest front load of any fund was 8.5% while the lowest was 1.5% with majority of front-load funds having loads between 4.50% and 6.49%. Third, there were 31 funds with deferred loads (for the five-year holding period).²² Interestingly enough, all of these 31 funds had a 1% deferred load given this five-year holding period and only one of these funds also had a front load. Fourth, there are many more growth and growth-income funds in the sample than there are aggressive growth, equity-income and small company funds. Fifth, young funds (funds with more than three but fewer than five years of return history) make up the smallest share of the sample. Middle-aged funds (funds with more than five but fewer than 10 years of return history) and seasoned funds (funds with 10 or more years of return history) together make up about 86% of the sample.

It should also be noted here that of the 304 no-load funds shown in Table 1, ten of these funds were index funds. This is important to note since the presence of index funds could affect our final results as they performed quite well over the out-of-sample period of 1993–1997. In most of the upcoming results, in addition to examining the full sample of funds, we also examine the results in which we exclude these 10 no-load index funds from the sample.²³

4.2. Overall results

Table 3 shows the overall out-of-sample results for all eight performance metrics. The results indicate that when using non-load-adjusted returns, there is no significant difference in average performance between no-load and load funds when using the

²¹ The timing of the mergers and liquidations was somewhat biased towards the first half of the out-of-sample period. Indeed, 60 of the merger/liquidations took place before July 1995 (the middle point of the out-of-sample period) and 37 occurred after this time.

²² 58 of the 635 funds had deferred loads but only 31 of the 58 had deferred loads that were still applied given the five-year holding period.

²³ There was also one load fund that was an index fund, however we do not exclude this fund from our sample as it does not affect the results at all.

Table 1
Summary statistics of the sample of funds

Samples organized by	No. of funds	Average front load (%)	Average deferred load (%)	No. of funds that liquidated between 1993 and 1997	No. of funds that merged between 1993 and 1997
<i>Overall</i>					
No-load funds	304	NA	NA	8	24
Load funds	331	See below	See below	12	53
<i>Load funds</i>					
Front-load funds	300	5.15	NA	10	45
Deferred-load funds	30	NA	1.00	2	7
Funds with front and deferred loads	1 ^a	2.50	1.00	0	1
<i>Style of fund</i>					
<i>Aggressive growth</i>					
No-load funds	17	NA	NA	0	1
Front-load funds	26	4.67	NA	2	2
Deferred-load funds	1	NA	1.00	0	0
<i>Equity income</i>					
No-load funds	22	NA	NA	1	1
Front-load funds	20	5.48	NA	0	3
Deferred-load funds	3	NA	1.00	0	2
<i>Growth</i>					
No-load funds	141	NA	NA	6	14
Front-load funds	149	5.16	NA	4	26
Deferred-load funds	12	NA	1.00	0	4
<i>Growth-income</i>					
No-load funds	73	NA	NA	1	6
Front-load funds	82	5.17	NA	3	11
Deferred-load funds	9	NA	1.00	2	1

Small company	No-load funds	51	NA	NA	0	2
	Front-load funds	23	5.30	NA	1	3
	Deferred-load funds	5	NA	1.00	0	0
	Funds with front and deferred loads	1	2.50	1.00	0	1
<i>Age of fund</i> Young (three to less than five years)	No-load funds	48	NA	NA	3	4
	Front-load funds	41	4.77	NA	3	7
	Deferred-load funds	1	NA	1.00	0	0
Middle-aged (five years to less than 10 years)	No-load funds	138	NA	NA	5	13
	Front-load funds	114	4.82	NA	7	18
	Deferred-load funds	24	NA	1.00	2	7
Seasoned (10 years or greater)	No-load funds	118	NA	NA	0	7
	Front-load funds	145	5.52	NA	0	20
	Deferred-load funds	5	NA	1.00	0	0
	Funds with front and deferred loads	1	2.50	1.00	0	1

The total number of funds is 635 funds. Loads, styles and ages of funds are as of December 31, 1992. Note that a deferred-load fund is a fund that still applied a deferred load after the investor had held the fund for five years. Note also that funds with front and or deferred loads are considered to be load funds.

^aThe one fund that had a front and deferred load was a seasoned small company fund.

Table 2
Summary statistics organized by the size of the load

Load of the fund	No. of funds	No. of funds that were liquidated between 1993 and 1997	No. of funds that merged between 1993 and 1997
<i>Front-load funds</i>			
Funds with front loads from 1.50% to 2.49%	5	0	0
Funds with front loads from 2.50% to 3.49%	15	0	3
Funds with front loads from 3.50% to 4.49%	29	4	4
Funds with front loads from 4.50% to 5.49%	120	4	18
Funds with front loads from 5.50% to 6.49%	99	2	16
Funds with front loads from 6.50% to 7.49%	12	0	2
Funds with front loads from 7.50% to 8.50%	20	0	2
<i>Total front-load funds</i>	300		
<i>Deferred-load funds</i>			
Funds with deferred loads (assuming the investor held the fund for five years) of 1%	30	2	7
<i>Total deferred-load funds</i>	30		
Funds with front and deferred loads (2.5% front-load and a 1% deferred load)	1	0	1

Note that a deferred-load fund is a fund that still applied a deferred load after the investor had held the fund for five years. Also note that the loads are defined as of December 31, 1992.

mean monthly returns and the Sharpe ratio. On the other hand, the results for the modified Jensen's alpha and the 4-index alpha indicate that the performance of no-load funds is significantly *higher* than that of load funds *even before adjusting* for loads in the returns (although in the case of the 4-index alpha significance is only at the 10% level). When using load-adjusted returns, all four of the performance metrics show significantly higher performance for the no-load funds (at the 1% level). Hence, these results show that the average performance of load funds is somewhat similar to that of no-load funds before adjusting for loads, however, after adjusting for loads, the average performance of load funds is far lower than that of no-load funds.

Table 3 also shows the average performance metrics when excluding the 10 no-load index funds. The results of this analysis are very similar to those where the index funds are included. Indeed, using the non-load-adjusted returns, the results for mean monthly returns and the Sharpe ratio show no significant difference between the load and no-load funds. However, the Jensen and 4-index alpha results indicate that the performance of no-load funds is significantly higher than the load funds. Using the load-adjusted returns, we again see that the results of all four performance metrics indicate that the no-load funds perform significantly better than the load funds at the 1% level of significance.

Table 3
Overall out-of-sample performance statistics^a

	No. of funds	Average mean monthly return (using non-load-adjusted returns)	Average monthly return (using load-adjusted returns)	Average Sharpe ratio (using non-load-adjusted returns)	Average Sharpe ratio (using load-adjusted returns)	Average Jensen alpha (using non-load-adjusted returns)	Average Jensen alpha (using load-adjusted returns)	Average 4-index alpha (using non-load-adjusted returns)	Average 4-index alpha (using load-adjusted returns)
No-load funds	304	0.9357	0.9357***	0.3025	0.3025***	-0.1612**	-0.1612***	0.0003*	0.0003***
Load funds	331	0.9295	0.8373	0.2931	0.2640	-0.2232	-0.3138	-0.0321	-0.1193
<i>Excluding index funds (excludes 10 no-load index funds from the sample)</i>									
No-load funds	294	0.9306	0.9306***	0.3007	0.3007***	-0.1637**	-0.1637***	0.0004*	0.0004***
Load Funds	301	0.9295	0.8373	0.2931	0.2640	-0.2232	-0.3138	-0.0321	-0.1193

* Indicates the difference between the no-load funds and load funds is significant at the 10% level.

** Indicates the difference between the no-load funds and load funds is significant at the 5% level.

*** Indicates the difference between the no-load funds and load funds is significant at the 1% level.

^a The out-of-sample performance statistics are for the five-year period from 1993 to 1997. Funds with front and or deferred loads are considered to be load funds.

4.3. Style effects

Table 4 presents the average performance results organized by style of fund using the styles as defined at the beginning of the out-of-sample period, December 31, 1992. There are two basic findings. First, unless you have a large enough sample of funds, it is difficult to see any strong pattern of significant differences in performance between no-load and load funds. This is evident in the findings for the aggressive growth and equity-income funds and to a lesser extent, for small company funds. Indeed for the aggressive growth and equity-income funds no significant differences are found and for the small company funds we find that two of the performance metrics (the load-adjusted Sharpe ratio and load-adjusted Jensen index) show that no-load funds significantly outperform load funds while one metric (the non-load-adjusted 4-index alpha) shows the opposite finding. Second, the patterns seen in Table 3 (the overall analysis) are again evident with respect to the growth and growth-income funds. That is, with non-load-adjusted returns performance measures we do not generally see much of a difference in performance between load and no-load funds. However, when using the load-adjusted returns all four of the performance metrics show that the no-load funds have significantly higher performance (at the 1% level) than do the load funds.

The bottom of Table 4 also shows the results from excluding the 10 no-load index funds (seven index funds were growth-income funds and three were small company). As with overall results in Table 3, the results are very similar to results on the full sample of funds.

In addition to examining the style effects using the styles as defined on December 31, 1992, we also examined the results using the styles of the funds *at the end of the sample period*, December 31, 1997.²⁴ The results of this analysis are presented in Table 5 and are similar to those presented in Table 4. The only major difference between Tables 4 and 5 is with the equity-income results. Table 4 shows that the difference in performance between the equity-income no-load and equity-income load funds is rarely significant, while Table 5 shows that there are three cases where that equity-income no-load funds significantly outperform equity-income load funds.

4.4. Age effects

Table 6 examines the effect that the age of the fund has on out-of-sample performance. As with the previous tables we examine the full sample of funds and a sample in which we exclude the no-load index funds.²⁵ We find two results. First, in general,

²⁴ The only exception is for merger and liquidated funds. For these funds we use the style of the fund defined during the quarter before they merged or liquidated. Note that none of the merger or liquidated funds changed their style from the beginning of the sample to the quarter before they merged or liquidated. Also note that 22 of the surviving funds changed styles outside of the five styles (aggressive growth, equity-income, growth, growth-income, small company) from the beginning of the sample to the quarter before they merged. We excluded these funds from the sample used in Table 5. Hence the sample consists of 613 funds.

²⁵ The index funds that were excluded included two young funds, six middle-aged funds and two seasoned funds.

Table 4
Out-sample performance statistics organized by style of fund (using style described by Morningstar as of December 31, 1992)^a

Fund style	Type of fund examined	No. of funds	Average excess monthly return (%) using non-load adjusted returns	Average excess monthly return (%) using load-adjusted returns	Average Sharpe ratio using non-load adjusted returns	Average Sharpe ratio using load adjusted returns	Average Jensen alpha using non-load adjusted returns	Average Jensen alpha using load adjusted returns	Average 4-index alpha using non-load adjusted returns	Average 4-index alpha using load adjusted returns	
Agg. growth	No-load	17	0.7420	0.7420	0.2037	0.2037	-0.4711	-0.4711	-0.0461	-0.0461	
	Load	27	0.8611	0.7734	0.2054	0.1850	-0.4764	-0.5625	-0.0089	-0.0919	
Equity-income	No-load	22	0.9222	0.9222	0.3800	0.3800*	0.0198	0.0198	-0.0620	-0.0620	
	Load	23	0.9714	0.8734	0.3772	0.3389	0.0469	-0.0495	-0.0184	-0.1110	
Growth	No-load	141	0.9372	0.9372***	0.2951	0.2951***	-0.1840**	-0.1840***	-0.0249*	-0.0249***	
	Load	161	0.9028	0.8097	0.2786	0.2500	-0.2789	-0.3705	-0.0752	-0.1633	
Growth-income	No-load	73	0.9810	0.9810***	0.3474	0.3474***	-0.0587	-0.0587***	-0.0375	-0.0375***	
	Load	91	0.9578	0.8649	0.3426	0.3090	-0.0909	-0.1822	-0.0640	-0.1519	
Small company	No-load	51	0.9370	0.9370	0.2582	0.2582**	-0.2195	-0.2195**	0.1665*	0.1665	
	Load	29	1.0195	0.9352	0.2332	0.2140	-0.3077	-0.3905	0.2745	0.1948	
<i>Excluding index funds</i>											
Growth-income	No-load	66	0.9657	0.9657**	0.3434	0.3434***	-0.0602	-0.0602***	-0.0376	-0.0376***	
	Load	91	0.9578	0.8649	0.3426	0.3090	-0.0909	-0.1822	-0.0640	-0.1519	
Small company	No-load	48	0.9334	0.9334	0.2564	0.2564*	-0.2218	-0.2218**	0.1717*	0.1717	
	Load	29	1.0195	0.9352	0.2332	0.2140	-0.3077	-0.3905	0.2745	0.1948	

* Indicates the difference between the no-load funds and load funds is significant at the 10% level.

** Indicates the difference between the no-load funds and load funds is significant at the 5% level.

*** Indicates the difference between the no-load funds and load funds is significant at the 1% level.

^a Note that the out-of-sample performance statistics are for the five-year period from 1993 to 1997. Also note that funds with front and deferred loads are considered load funds.

Table 5
Out-sample performance statistics organized by style of fund (using styles described by Morningstar as of December 31, 1997)^a

Fund style	Type of fund examined	No. of funds	Average excess mean monthly return (%)		Average Sharpe ratio		Average Jensen alpha		Average alpha		Average 4-index alpha	
			using non-load adjusted returns	load-adjusted returns	using non-load adjusted returns	load-adjusted returns	using non-load adjusted returns	load-adjusted returns	using non-load adjusted returns	load-adjusted returns		
Agg. growth	No-load	18	0.7859	0.7859	0.1791	0.1791	-0.4747	-0.4747	-0.0198	-0.0198	-0.0198	-0.0198
	Load	24	0.8393	0.7527	0.1986	0.1787	-0.5964	-0.5113	-0.0921	-0.0102	-0.0921	-0.0921
Equity-income	No-load	18	0.9724	0.9724**	0.3919	0.3919**	0.0276*	0.0276	-0.0727	-0.0727	-0.0727	-0.0727
	Load	25	0.9656	0.8693	0.3758	0.3381	0.0404	0.0404	-0.1147	-0.0236	-0.1147	-0.1147
Growth	No-load	125	0.9510	0.9510***	0.2979**	0.2979***	-0.2001***	-0.2001***	-0.0242**	-0.0242*	-0.0242***	-0.0242***
	Load	157	0.9013	0.8079	0.2722	0.2440	-0.2935	-0.2935	-0.1611	-0.0727	-0.1611	-0.1611
Growth-income	No-load	75	1.0093	1.0093***	0.3610	0.3610***	-0.0346**	-0.0346**	-0.0267	-0.0267	-0.0267***	-0.0267***
	Load	86	0.9875	0.8959	0.3485	0.3161	-0.0850	-0.0850	-0.1475	-0.0608	-0.1475	-0.1475
Small-company	No-load	55	0.9227*	0.9227	0.2558	0.2558*	-0.2185	-0.2185	0.1380	0.1380*	0.1380	0.1380
	Load	30	1.0140	0.9273	0.2403	0.2197	-0.2849	-0.2849	0.1683	0.2503	0.1683	0.1683
<i>Excluding index funds</i>												
Growth-income	No-load	68	0.9974	0.9974**	0.3585	0.3585***	-0.0335***	-0.0335***	-0.0257	-0.0257	-0.0257***	-0.0257***
	Load	86	0.9875	0.8959	0.3485	0.3161	-0.0850	-0.0850	-0.1475	-0.0608	-0.1475	-0.1475
Small-company	No-load	52	0.9185*	0.9185	0.2540	0.2540	-0.2206	-0.2206	0.1411*	0.1411*	0.1411	0.1411
	Load	30	1.0140	0.9273	0.2403	0.2197	-0.2849	-0.2849	0.1683	0.2503	0.1683	0.1683

* Indicates the difference between the no-load funds and load funds is significant at the 10% level.

** Indicates the difference between the no-load funds and load funds is significant at the 5% level.

*** Indicates the difference between the no-load funds and load funds is significant at the 1% level.

^a The styles are the fund styles as described by Morningstar on December 31, 1997, unless the fund was a merger or liquidated fund in which case we used the style as defined on the quarter before the fund merged or liquidated. Note that 22 funds that were classified as either aggressive growth, equity-income, growth, growth-income or small company in December 1992, were classified as something other than one of these styles as of December 31, 1997. Hence these funds were excluded from the sample. Consequently the sample size was 613 funds. Note also that the out-of-sample performance statistics are for the five-year period from 1993 to 1997 and that front-load and deferred-load funds are considered load funds.

Table 6
Age, loads and out-of-sample performance^a

Age of fund	Fund type	No. of funds	Average Sharpe ratio (using non-load adjusted returns)	Average Sharpe ratio (using load adjusted returns)	Average Jensen alpha (using non-load adjusted returns)	Average Jensen alpha (using load adjusted returns)	Average 4-index alpha (using non-load adjusted returns)	Average 4-index alpha (using load adjusted returns)
Young	No-load funds	48	0.3342	0.3342***	-0.0784	-0.0784***	0.0182	0.0182**
	Load funds	42	0.3045	0.2750	-0.1383	-0.2275	-0.0074	-0.0931
Middle-aged	No-load funds	138	0.3082*	0.3082***	-0.1535***	-0.1535***	0.0151	0.0151***
	Load funds	138	0.2858	0.2602	-0.2447	-0.3255	-0.0211	-0.0989
Seasoned	No-load funds	118	0.2829	0.2829	-0.2038	-0.2038***	-0.0243	-0.0243***
	Load funds	151	0.2967	0.2643	-0.2272	-0.3272	-0.0490	-0.1453
<i>Excluding index funds</i>								
Young	No-load funds	46	0.3318	0.3318**	-0.0806	-0.0806**	0.0203	0.0203**
	Load funds	42	0.3045	0.2750	-0.1383	-0.2275	-0.0074	-0.0931
Middle-aged	No-load funds	132	0.3064*	0.3064***	-0.1554**	-0.1554***	0.0159	0.0159***
	Load funds	138	0.2858	0.2602	-0.2447	-0.3255	-0.0211	-0.0989
Seasoned	No-load funds	116	0.2819	0.2819	-0.2062	-0.2062***	-0.0253	-0.0253***
	Load funds	151	0.2967	0.2643	-0.2272	-0.3272	-0.0490	-0.1453

* Indicates the difference between the no-load funds and load funds is significant at the 10% level.

** Indicates the difference between the no-load funds and load funds is significant at the 5% level.

*** Indicates the difference between the no-load funds and load funds is significant at the 1% level.

^a Note that young funds have at least three but less than five years of return history as of December 31, 1992; middle-aged funds have five years to less than ten years of returns as of December 31, 1992; and seasoned funds have ten years or more of returns as of December 31, 1992. Also note that funds with front and deferred loads are considered load funds. The out-of-sample performance statistics are for the five-year period from 1993 to 1997.

we find the same results as those in Table 3. Across different ages of funds, there is little difference in the performance of load and no-load funds when using non-load-adjusted performance metrics. Indeed, across the three performance metrics²⁶ there are only two cases (Sharpe ratio and Jensen alpha for middle-aged funds) where there is a significant difference between load and no-load funds when using non-load-adjusted returns. However, when the load-adjusted performance metrics are used, we find strong evidence that no-load funds significantly outperform the load funds.

Second, we find that the young funds dominate in terms of performance. In every single performance metric, whether load adjusted or not, the young funds have the highest average performance metrics. These results are consistent with others (e.g. Chevalier and Ellison, 1997) who find that young funds significantly outperform older funds.

4.5. Size of the load effect

Another issue we investigate is whether the size of the load helps explain the difference in performance between no-load and load funds. To investigate this issue we examine differences in average performance between no-load funds and load funds organized by the size of the load. To do this we classify the load funds into four groups: *deferred* load funds (which all have deferred loads of 1%), *low front* load funds²⁷ (funds with front loads of 4% or less), *middle front* load funds (funds with front-loads between 4.01% and 5.99%) and *high front* load funds (funds with front-loads of 6% or more).²⁸ These results are presented in Table 7.²⁹

We find that in most cases (all but two cases) the raw differences indicate that no-load funds outperform the load funds, however the difference is only significant in about half the cases. More specifically, the results show that only four of the 12 differences between no-load funds and low-load funds (deferred and low-front load funds) are in fact significant. On the other hand, when examining the difference between no-load and higher-load funds (middle-front load and high-front load funds), the results show much stronger evidence that no-load funds significantly outperform load funds. Indeed, we find that the difference between no-load funds and middle-load funds and the difference between no-load and high-load funds are significant across all three performance metrics when using load-adjusted performance. Hence, although the results show more significant differences between no-load and high-load funds, the fact that the differences generally indicate that no-load fund perform

²⁶ We did not use mean monthly returns since the use of different styles would yield inconsistent results.

²⁷ Note that the one fund which had both a deferred load and a front-load was classified as a low-front load fund.

²⁸ Note that other breakdowns were used, however in no instance was there a substantial change in the results.

²⁹ Note that the results reported are those in which we excluded the ten no-load index funds. The results were very similar when we included the index funds. Also note that we did not use mean monthly returns since the use of different styles would yield inconsistent results.

Table 7
Difference in average performance between no-load and load funds organized by size of load^a

	No. of load funds	Average load (%)	Non-load adjusted Sharpe ratio	Load-adjusted Sharpe ratio	Non-load adjusted Jensen alpha	Load-adjusted Jensen alpha	Non-load adjusted 4-index alpha	Load adjusted 4-index alpha
Difference between no-load and deferred-load funds	30	1.0	0.0146	0.0240	0.0704	0.0988*	0.0128	0.0400
Difference between no-load and low-front-load funds	50	3.38	-0.0034	0.0191	0.0172	0.0844*	0.0623*	0.1269***
Difference between no-load and middle-front-load funds	214	5.16	0.0124	0.0428***	0.0755**	0.1725***	0.0230	0.1165***
Difference between no-load and high-front-load funds	37	7.47	-0.0110	0.0358**	0.0155	0.1513***	0.0631*	0.1938***

Note: *Deferred-load funds*: 1% loads (assuming the investor held onto the fund for the five-year holding period) as of December 31, 1992. *Low-front load funds*: 1.5–4.0% loads as of December 31, 1992. *Middle-front-load funds*: 4.01–5.99% loads as of December 31, 1992. *High-front-load funds*: 6.0–8.5% loads as of December 31, 1992.

* Indicates the difference between the no-load funds and load funds is significant at the 10% level.

** Indicates the difference between the no-load funds and load funds is significant at the 5% level.

*** Indicates the difference between the no-load funds and load funds is significant at the 1% level.

^a Note that for no-load fund performance calculations we *excluded* the 10 no-load index funds. All the out-of-sample performance statistics are for the five-year period from 1993 to 1997. Also note that the one fund that had both a front and deferred load was considered to be a low-front load fund.

Table 8
Load size and out-of-sample performance: A regression analysis^a

Out-of-sample performance metric	γ_0	γ_1
Non-load adjusted Sharpe ratio	0.2801***	0.0027
Load-adjusted Sharpe ratio	0.2768***	-0.0027
Non-load adjusted jensen alpha	-0.2428***	0.0041
Load-adjusted jensen alpha	-0.2562***	-0.0121
Non-load adjusted 4-index alpha	-0.0229	-0.0019
Load-adjusted 4-index alpha	-0.0356	-0.0176**

Note also that for the one fund that had a deferred and front load we converted the deferred load to a front load and added this to the existing front load.

** Significant from zero at the 5% level.

*** Significant from zero at the 1% level.

^a The following are the results of the equation: $S_i = \gamma_0 + \gamma_1(\text{load})_i + u_i$, where S_i is the out-of-sample performance metric value for the period 1993–1997, and load_i is the front load of the fund or the deferred load (for the five year holding period) converted into a front load. Hence load_i only measures the front load. All loads are found from the December 31, 1992 data disk. Note also that only load funds (both deferred and front-load funds) are examined. Hence, the sample of each regression is 331 funds.

better than load funds indicates that even across the size of the load, no-load funds generally outperform load funds.

4.6. Size of load effect in load funds

One last issue we examine is whether the size of the load influences the performance of load funds. This issue is particularly relevant for the investor who has decided to buy a load fund but is using the size of the load to determine the selection. To investigate this issue we examined the following equation:

$$S_i = \gamma_0 + \gamma_1 \text{load}_i + u_i \quad (5)$$

where S_i is the risk-adjusted out-of-sample performance metric value for the period 1993–1997 for the i th fund. load_i is the front load of the fund or the deferred load of the fund (for the five-year holding period) converted into a front load.³⁰ In the regression, only load funds are examined. Hence, the sample size of each regression was 331 funds (300 front-load funds, 30 deferred load funds, and one fund with a front and deferred load).³¹ The results are presented in Table 8.

We find, surprisingly, that there is little evidence of a statistically significant relationship between the out-of-sample performance metric and the level of the front-load. Indeed, the γ_1 terms for all the performance metrics except for the load-adjusted 4-index alpha are all close to zero and clearly not significant. Although

³⁰ In order to include deferred-load funds into the regression we had to normalize the deferred loads into front loads. To do this we converted each deferred load (for the five-year holding period) into its equivalent front load. More specifically, we assumed that instead of pre-paying the deferred load as indicated in Eq. (2), we assume that the investor borrows the amount necessary. A 1% deferred load was thus a 0.7427% front load.

³¹ For the one fund that had a deferred and a front load we combined the front load with the converted deferred load to get the load for the fund.

Table 9
Load size and out-of-sample performance: Average out-of-sample performance organized by the size of the load^a

Fund type	No. of load funds	Average load (%)	Average Sharpe ratio (using load adjusted returns)	Average Sharpe ratio (using non-load adjusted returns)	Average Jensen alpha (using non-load adjusted returns)	Average Jensen alpha (using load adjusted returns)	Average 4-index alpha (using non-load adjusted returns)	Average 4-index alpha (using load adjusted returns)
Deferred-load funds	30	1.0	0.2860	0.2767	-0.2341	-0.2625	-0.0124	-0.0396 ^b
Low-front-load funds:	50	3.38	0.3041	0.2816	-0.1809	-0.2481	-0.0619	-0.1265
Funds with loads of 1.5–4%								
Middle-front-load funds: Funds with loads of 4.01–5.99%	214	5.16	0.2883	0.2579	-0.2392	-0.3362	-0.0226	-0.1161 ^b
High-front-load funds: Funds with loads of 6.0–8.5%	37	7.47	0.3117	0.2649	-0.1792	-0.3150	-0.0627	-0.1934 ^b

Note that one front load fund also had a deferred load. Hence the returns for this one fund are front and deferred load adjusted. This fund was part of the low-front-load funds group.

^aThe 331 load funds¹ (as of December 31, 1992) are rank ordered into four groups by their load as of December 31, 1992. The out-of-sample performance of each of the four groups is then examined. The out-of-sample performance period is 1993–1997.

^bWe tested to see if there were any significant differences between the types of load funds. There was a significant difference found (at the 5% level) between the load-adjusted 4-index alpha of high-front-load funds and the load-adjusted 4-index alpha of deferred load funds. There was also a significant difference found (at the 5% level) between the load-adjusted 4-index alpha of high-front load funds and the load-adjusted 4-index alpha of middle-front-load funds. There were no other significant differences found.

insignificant, the signs on the γ_1 coefficients illustrate that when using non-load-adjusted returns, there is a tendency for higher-load funds to have *higher* out-of-sample performance. Conversely, when using load-adjusted returns, the signs are reversed indicating that the higher the load the lower the out-of-sample performance.

To further illustrate our results we present Table 9 which shows the average performance metrics for deferred-load funds, low-front load funds, middle-front load funds and high-front load funds. The results here show similar overall findings to those of Table 8: high-load funds perform slightly better than low-load funds before adjusting for loads, however, this performance difference reverses itself when loads are factored into the returns.

The only exception to this conclusion from Table 9 concern the results for the 4-index alpha which shows that the deferred load funds (hence very low load funds) actually have higher performance for *both* the non-load and load-adjusted returns. Furthermore the results for the load-adjusted 4-index alpha metric also show that deferred load funds significantly outperform high-load funds and that middle-front funds load significantly outperform high-load funds.

Nevertheless, the general picture of these results suggest that if an investor wants to hold a load fund, the size of the load is not a strong sign of future non-load or load-adjusted performance. Before loads are assessed, high-load funds may perform just slightly better than low-load funds, yet that difference is not enough to compensate for the higher loads.

5. Conclusions

In light of a recent industry trend towards load funds and away from no-load funds, this paper has examined and compared the out-of-sample performance of no-load and load mutual funds. Unlike the previous literature on this topic, this paper provides a more comprehensive analysis as it uses methodologies to incorporate loads directly into the returns, utilizes a large sample of funds free of survivorship bias, and evaluates performance across many different performance metrics and different ages and styles of funds. We find two important results.

1. Before adjusting for loads in the returns, no-load funds perform somewhat better than load funds although the differences are often not significant. However, after adjusting for loads in the return data, no-load funds are found to perform much better than do load funds, with the differences found to be significant at the 1% level across many different performance metrics. This difference is also found when funds are disaggregated by age, style and to some degree, by size of the load. Since this paper is first to incorporate loads directly into the performance numbers, it provides the first quantifiable indication of the difference between investor-realized return performance of load and no-load funds. Our findings indicate that investors are sacrificing a significant degree of performance to hold load funds. Indeed, the load adjustment that is used here is calculated during a time of low interest rates (1993). If in fact interest rates had been higher at the time the investor bought a load fund, the loss in performance arising from the load adjustment would have even been more severe.

2. Although we document a substantial difference in performance between no-load and load funds, we also find that among load funds there is little significant difference in out-of-sample performance between high-load funds and low-load funds even after adjusting for loads. Indeed, high-load funds only perform slightly better than low-load funds before loads are assessed and only slightly worse after loads are assessed. Hence, this study documents that for the investor who prefers the services provided by load funds, the size of the load has little predictive ability in determining future fund performance. Of course, this paper has also shown that if performance is what investors are mainly concerned with, they should not consider load funds.

Acknowledgements

I would like to thank participants at the Scottish Institute of Finance (SIRF) conference on Managed Funds (Summer 2001) for helpful comments and to especially thank an anonymous referee for extremely valuable comments.

References

- Blume, M.E., 1998. An anatomy of Morningstar ratings. *Financial Analysts Journal* 54 (2), 19–27.
- Brown, S.J., Goetzmann, W.N., Ibbotson, R.G., Ross, S.A., 1992. Survivorship bias in performance studies. *Review of Financial Studies* 5, 553–580.
- Carlson, R.S., 1970. Aggregate performance of mutual funds 1948–1967. *Journal of Financial and Quantitative Analysis* 5 (1), 1–32.
- Chevalier, J., Ellison, G., 1997. Risk taking by mutual funds as a response to incentives. *Journal of Political Economy* 105 (6), 1167–1200.
- Detzler, M.L., 1999. The performance of global bond mutual funds. *Journal of Banking and Finance* 23 (8), 1195–1217.
- Droms, W.G., Walker, D.A., 1994. Investment performance of international mutual funds. *Journal of Financial Research* 18 (1), 1–14.
- Elton, E.J., Gruber, M.J., Blake, C.R., 1996a. The persistence of risk-adjusted mutual fund performance. *Journal of Business* 69 (2), 133–157.
- Elton, E.J., Gruber, M.J., Blake, C.R., 1996b. Survivorship bias and mutual fund performance. *Review of Financial Studies* 9 (4), 1097–1120.
- Elton, E.J., Gruber, M.J., Das, S., Hlavka, M., 1993. Efficiency with costly information: A reinterpretation of evidence from managed portfolios. *Review of Financial Studies* 6 (1), 1–22.
- Fortin, R., Michelson, S.E., 1995. Are load mutual funds worth the price? *Journal of Investing* 4 (3), 89–94.
- Grinblatt, M., Titman, S., 1994. A study of monthly mutual fund returns and performance evaluation techniques. *Journal of Financial and Quantitative Analysis* 29 (3), 419–444.
- Gruber, M.J., 1996. Another puzzle: The growth in actively managed mutual funds. *Journal of Finance* 51 (3), 783–810.
- Ippolito, R.A., 1989. Efficiency with costly information: A study of mutual fund performance 1965–1984. *Quarterly Journal of Economics* 104 (1), 1–23.
- Jensen, M., 1968. The performance of mutual funds in the period 1945–1964. *The Journal of Finance* 23, 389–416.
- Kihn, J., 1996. To load or not to load? A study of marketing and distribution charges of mutual funds. *Financial Analysts Journal* 52 (3), 28–36.
- Morningstar Principia Manual, 1998. Morningstar Publications, Chicago, IL.
- Rea, J.D., Reid, B.K., 1998. Trends in the ownership cost of equity mutual funds. *Investment Company Institute Perspective* 4 (3), 1–15.
- Sharpe, W., 1966. Mutual fund performance. *Journal of Business* 39 (1), 119–138.

74

NBER WORKING PAPER SERIES

THE MARKET FOR FINANCIAL ADVICE:
AN AUDIT STUDY

Sendhil Mullainathan
Markus Noeth
Antoinette Schoar

Working Paper 17929
<http://www.nber.org/papers/w17929>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2012

We thank Tim Adam, John Campbell, Matthew Gentzkow, Michael Haliassos, Steffen Meyer, Jay Ritter, and seminar participants at Berkeley, Columbia, MIT Sloan, NBER Behavioral Economics Working Group, NBER Household Finance workshop, NetSpar Torino, SAVE Deidesheim, and ESSFM Gerzensee as well as at the AEA and the AFA for their helpful comments. We thank Ximena Cadena and Carter Powers for outstanding research assistance in implementing the project. We would also like to acknowledge financial support for the audit study from ideas42. All mistakes are of course our own. Sendhil Mullainathan is the assistant director of research at the CFPB. The views expressed are those of the authors and do not necessarily represent those of the Director of the Consumer Financial Protection Bureau nor those of the staff, nor the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2012 by Sendhil Mullainathan, Markus Noeth, and Antoinette Schoar. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Market for Financial Advice: An Audit Study
Sendhil Mullainathan, Markus Noeth, and Antoinette Schoar
NBER Working Paper No. 17929
March 2012
JEL No. G02,G1,G11,G2,G23,G24

ABSTRACT

Do financial advisers undo or reinforce the behavioral biases and misconceptions of their clients? We use an audit methodology where trained auditors meet with financial advisers and present different types of portfolios. These portfolios reflect either biases that are in line with the financial interests of the advisers (e.g., returns-chasing portfolio) or run counter to their interests (e.g., a portfolio with company stock or very low-fee index funds). We document that advisers fail to de-bias their clients and often reinforce biases that are in their interests. Advisers encourage returns-chasing behavior and push for actively managed funds that have higher fees, even if the client starts with a well-diversified, low-fee portfolio.

Sendhil Mullainathan
Department of Economics
Littauer M-18
Harvard University
Cambridge, MA 02138
and Consumer Financial Protection Bureau
and also NBER
mullain@fas.harvard.edu

Antoinette Schoar
MIT Sloan School of Management
100 Main Street, E62-638
Cambridge, MA 02142
and NBER
aschoar@mit.edu

Markus Noeth
LS Banking & Behavioral Finance
VMP 5
University of Hamburg
20146 Hamburg
Germany
markus.noeth@uni-hamburg.de

1. Introduction

A large and still growing body of research demonstrates that individual investors often make poor financial decisions if left to their own devices. Drawing on evidence from psychology and behavioral economics, these studies suggest that investor beliefs and decision processes are prone to biases that often result in financial decisions at odds with basic portfolio theory. For example, retail investors are overconfident, engage in trend-chasing, use naïve heuristics, and are generally susceptible to a number of different biases. Benartzi and Thaler (2007) document some of the most common biases.¹ Using trade data from retail customers, Barber and Odean (2000) document that excess trading in brokerage accounts and returns-chasing behavior leads to significantly lower returns compared to a buy-and-hold strategy. In short, the past decade of research has produced a large body of evidence suggesting that households may be bad at choosing portfolios on their own.²

Yet households do not make decisions in a vacuum. A variety of forces, from social interactions with friends and family to advertising and media, can influence their choices. One particularly important source of inputs comes from financial advisers.³ In a survey of retail investors, Hung et. al. (2008) found that 73% of all individuals surveyed consult a financial adviser before purchasing shares or mutual funds. Given this central role of advisers in the investment process, we ask whether or not the market for financial advice serves to de-bias individual investors and thus correct potential mistakes they might make without these inputs.⁴ We define ‘good advice’ as advice that moves the investor toward a low-cost, diversified, index-fund approach, which many textbook analyses on mutual fund investments suggest, see for example Carhart (1997). Alternatively, since many (though not all) advisers are paid with incentives that encourage them to direct money to specific funds and generate high fees, might advisers exploit these biases of retail investors in order to further their own interests? Additionally, if investors are unable to make portfolio decisions that are in line with text book recommendations, they might be equally unable to differentiate between advisers that are either self-interested or help investors to build a well diversified, low fee portfolio. In such a world, good advice, i.e. , to hold a well-diversified portfolio composed of low-fee funds, might not be rewarded sufficiently.⁵ As a

¹ Benartzi and Thaler (2001) for example argue that employees follow a naïve diversification strategy of mechanically spreading their money equally across the funds they are offered (what they call 1/n rule), generating quite perverse outcomes since the equity mix depends on the investment menu.

² It is not in the scope of this article to review this body of literature. See Barberis and Thaler (2003) and Campbell (2006) for additional in depth overviews.

³ Most financial advisers provide personalized advice of what stocks or funds to invest in. The Investment Advisers Act of 1940 defines (see section 202(11)): “Investment adviser’ means any person who, for compensation, engages in the business of advising others, either directly or through publications or writings, as to the value of securities or as to the advisability of investing in, purchasing, or selling securities, or who, for compensation and as part of a regular business, issues or promulgates analyses or reports concerning securities; (...).” Bolton, Freixas, and Shapiro (2007) study competition between banks and its influence on incentives for truthful information revelation.

⁴ Chen, De, Hu, and Hwang (2011) focus on the impact of peer-based financial advice via social media on aggregate stock market outcomes.

⁵ Of course, understanding this effect requires a deeper understanding of a different source of cognitive bias-people’s perceptions of what constitutes good advice and what constitutes independent advice. See Moore, Cain, Loewenstein and Bazerman (2005) for work on this topic. In this context, it is also an open question why retail investors are mostly unaware of a conflict of interest between the provider of information and its recipient. Malmendier and Shanthikumar (2007) show that retail investors are mostly naïve with respect to analysts’ recommendation incentives. In addition, Chatter, Inderst, and Huck (2010) document in a

consequence, these advisers might be driven out of the market in equilibrium, since they cannot attract a sufficient market share to account for their lower-fee structure. The experience of Vanguard, which offered the first index funds, is a case in point: the firm had to modify their sole reliance on low-cost index funds since consumers seemed to be susceptible to high-cost advertising (see Bogle, 2000). Thus, the market for financial advice may not mitigate behavioral biases and may even exacerbate them. Despite the (growing) importance of advisers in the investment process, especially due to an increasing number of defined contribution plans, very little is known about the market for financial advice.⁶ Campbell (2006) highlights our relative ignorance about this important sector.

Understanding this market is also important from an aggregate level. The market for financial advice might influence how individual biases translate into aggregate market outcomes, e.g., capital flows into different investment strategies or even pricing. How well it de-biases individuals is important for knowing how to model “representative” agents in macro-consumption models and how to model equity prices and for numerous policy applications.

In this paper we set up an audit study to test the quality of advice provided to retail investors in the market. The specific advisers we are looking at in this study are retail advisers whom the average citizen can access via their bank, independent brokerages, or investment advisory firms. These advisers are usually paid based on the fees they generate but not based on the assets under management or the performance of the portfolio. Once clients have more than US\$ 500k in investible capital, they have access to a broader set of advisers with better compensation structures. We focus on pure investment advisers in order to focus on a narrow set of measureable outcomes. Therefore, we do not include tax advisers, advisers who also provide estate planning services, or providers of other wealth management services.

In particular, we want to understand whether advisers actively de-bias their clients or instead exaggerate existing biases, especially biases that help the adviser’s own interests by increasing fees and turnover. For that purpose, auditors were randomly assigned to four different treatments (portfolios) which represent different investment strategies and biases. We make sure that the loss to the client from each of the biases is comparable. The random assignment allows us to test the average response of a typical adviser without any concern for self-selection of clients to different types of advisers. The auditors tracked the information requested from them, the advice given, and other features of the interaction. Our protocol records the advice given via auditor exit surveys, as well as written materials with portfolio suggestions by the adviser.

The first two portfolios (“bias scenarios”) reflect two of the most commonly described biases in the literature. In the first scenario (“chasing fund returns”), the auditor holds a portfolio in which 30% is invested in one sector exchange traded fund that performed well in the previous year, and he expresses an interest in identifying more industries that had done well recently. In this case, the incentives of the adviser and of the client are not aligned: the adviser benefits from the bias of the client since it allows him to churn the portfolio more often and generate more fees, whereas the client would profit from a better

European Survey with several thousand participants that retail investors ignore advisers’ potential conflict of interest. Inderst and Ottaviani (2011b) investigate theoretically the determinants of the compensation structure for brokers. Also see Schneider (2009) about trust and incentives structures.

⁶ The market for financial advice generates between US\$ 20bn and US\$ 50bn fees per year depending on the definition of advice and compensation models for advisers.

diversified portfolio.⁷ In the second scenario (“employer stocks”), an auditor holds 30% of his portfolio in the company stock of his assigned employer. Thus, incentives of the adviser and of the client are aligned: it is in the best interest of the adviser to reduce or eliminate the client’s bias since holding company stock also reduces the adviser’s ability to generate fees.⁸ In the third scenario, the auditor holds a diversified, low-fee portfolio consisting of index funds and bonds--in effect, an efficient US portfolio. We introduce this scenario to test if advisers are willing to move clients out of this portfolio which would be closest to an allocation recommended in most finance textbook. Finally we have a control group (“cash scenario”) in which the advisee simply holds certificates of deposits and does not espouse a particular view beyond a general willingness to increase risk for higher returns. This variation in treatment groups will allow us to test how responsive advisers are to the needs of their prospective clients.

Our audit produces three main sets of findings. First, advisers’ reactions to different portfolios or investment scenarios varied significantly. Advisers were broadly supportive of the trend-chasing portfolio but much less supportive of the company stock portfolio. Most strikingly, they were unsupportive of the (efficient) index portfolio and suggested a change to actively managed funds. Overall, advisers had a significant bias towards active management. In nearly 50% of the visits, the adviser encouraged investing in an actively managed fund; by contrast, in only 7.5% of the advice sessions (21 visits), advisers encouraged investing in an index fund. When advisers mentioned fees, they did so in a way that downplayed them without lying. For example, they often used arguments like, “This fund has 2% fee but that is not much above industry average.” These results suggest that the market for financial advice does not serve to de-bias clients but in fact exaggerates biases that are in the adviser’s financial interest while leaning against those that do not generate fees. In our index fund scenario, the advisers are even advocating a change in strategy (away from low fee index funds and towards high fee actively managed funds) that would make the client worse off than the allocation with which he or she started off.

Second, consistent with portfolio theory, most advisers did ask clients about their demographic characteristics, which may have been used to determine risk preferences, time-horizon and human capital risks, and covariance. Overall, we find that in more than 75% of the visits, advisers asked for this kind of information, specifically income, other savings (e.g. 401(k) plan) besides what they were investing with the adviser, occupation, and marital and parental status. The recommended investment in stocks and domestic assets significantly increased with annual income, a fact that may be explained by an assumed higher risk or loss tolerance for the well-off. Married clients were told to hold less in liquid assets. This is consistent with a model of spousal labor supply providing insurance, reducing the need for liquidity. However, in many cases, the information did not get used in the way predicted by portfolio theory: the recommended exposure to equities decreased with the amount invested. Female clients were asked to hold more liquidity, advised to hold less international exposure, and pushed less frequently to invest in actively managed funds. At the same time, advisers did not seem to tailor portfolio advice with the age of the client at hand. We find no significant differences in the mix of stocks and bonds for older clients. By and large, though, this is the arena where advisers were closest to traditional theory:

⁷ Inderst and Ottaviani (2009, 2011a) link advisers compensation and their advice quality theoretically either by focusing on the agency relationship between the selling firm and its employed sales force or by analyzing competition through commissions and kickbacks paid to advisers by the fund industry.

⁸ The advice to sell the employer stocks and to invest the money in a diversified portfolio enhances the client’s portfolio and generates fees for the adviser.

attempting to match portfolios to characteristics. The levels of portfolio advice were also broadly consistent with portfolio theory, with advisers suggesting a high equity mix (roughly 2/3) and thereby potentially reducing any bias that may generate an equity premium. Interestingly, they were also more likely to mention fees spontaneously when the auditor was older, possibly believing that older auditors would have asked themselves.⁹ Their responses to the different portfolios reinforce these facts.

Third, we find some suggestive evidence of ‘catering’, i.e. advisers showed support early on for the client’s existing strategies, most likely to establish credibility and not alienate a potential client. The “initial reaction” to a client’s strategy varied significantly from the later recommended course of action. In fact, the initial reaction to index fund investments was very positive while auditors who went in with company stock or returns-chasing portfolios faced more ambiguous support. These results highlight that advisers have to be aware of the fact that they are facing a sales situation and they cannot bluntly criticize what clients might have done in the past. However, it does not appear that advisers are severely limited in the investment recommendations they eventually give to their clients, since advisers have no problem discouraging clients from investing more in their current strategies (especially if it goes against the adviser’s interests). Instead, they suggest investments that are orthogonal to the client’s current approach.

Overall our findings suggest that the market for advice works very imperfectly. The advice by and large fails to de-bias clients and if anything may exaggerate existing biases or, in some cases, even makes the clients worse off. Moreover, individual biases can have first-order implications for aggregate capital flows and pricing of risk, if there is not enough informed capital to exploit arbitrage opportunities against capital flows from “biased” retail investors. It can also shed light on how we model information aggregation in equilibrium if competitive forces in the market for financial advice do not lead to the provision of the best possible advice. Competition might be limited by the fact that financial advisers exploit the biases of naïve (or uninformed) retail investors. At the same time, advisers who are interested in providing better advice might be unable to gain a market share if biased retail investors are unable to differentiate good from bad advice. While we cannot rule out the latter force, our evidence suggests that advisers’ self-interest plays an important role in providing advice that is not in the best interests of their clients.

It is important to note that our research design allows us to look at the quality of investment advice provided to clients. However, advisers may provide many other benefits for their clients, for example, by giving them the confidence and information to invest in risky assets in the first place, by protecting them from losing money in fraudulent funds, or by reducing transaction costs. These reasons might be as important as the actual content of the advice. However, even if these additional benefits are very important for retail investors (perhaps even the primary reason that clients seek advisers), there is no reason why advisers should not be able to also provide high-quality advice to their customers.

While audit studies have been used to measure discrimination in the labor and housing markets (Fix and Turner (1998), Altonji and Blank (1999) and Heckman (1998) for reviews), they have not been used in the financial context, save a few exceptions (see recent work by Iyer and Schoar, 2009 and 2011). Importantly, the audit study methodology allows us to measure an adviser’s response when we exogenously vary the types of clients and biases that the adviser is confronted with. This enables us to control for the selection of

⁹ Note that “mentioning fees” may include statements like “this is a no-load fund”, i.e. not all relevant fees are mentioned.

clients to advisers, which is one of the central problems plaguing non-experimental data. The reason is that non-experimental data (e.g., bank records) usually does not allow us to differentiate how much of the observed outcomes are driven by the adviser influencing the clients' decisions versus different clients selecting to certain types of advisers who will provide them with the recommendations they want to hear. In addition, the audit methodology enables us to vary the characteristics of the auditor exogenously either by selecting certain auditors or by assigning specific characteristics. In addition to demographic characteristics like age, gender, and number of children, we also varied traits such as investible wealth and housing status.

Our results are also related to a small but growing literature on financial advisers. One of the most noted early studies on financial advice is Canner, Mankiw, and Weil (1997), who examine the generic written advice given by investment advisers based on broad rules of thumb (see also Bodie and Crane, 1997). More recent work by Bergstresser, Chalmers, and Tufano (2009) or Del Guercio, Reuter, and Tkac (2011) show the role of incentives and distribution channels in the provision of financial advice. Bluethgen et al. (2008), Chalmers and Reuter (2011), Hackethal et al. (2011, 2012), Kramer (2012), and Bhattacharya et al. (2012) use data on portfolio outcomes or trading volume to quantify the benefits of financial advice. Georgarakos and Inderst (2011) show theoretically and provide some empirical evidence that trust in professional financial advice has a statistically and economically significant effect on the stock market participation for households with low financial capability, i.e., most retail investors. Similarly, Inderst, and Ottaviani (2009, 2011a) provide a theoretical framework that links adviser compensation with advice quality, focusing on the agency relationship between the advisory firm and its sales force and competition via commissions and kickbacks which are paid to advisers by the fund industry.

The rest of the paper proceeds as follows: section 2 provides a summary of the study design and setup of the audit. Section 3 shows the descriptive statistics and confirms the randomization. We report the main set of results in section 4 and conclude in section 5.

2. Study Design

2.1. Overview of Audit

In order to investigate the quality of financial advice that is commonly given to clients in the market, we set up an audit study in the greater Boston and Cambridge area. We sent trained, professional auditors to impersonate regular customers who are seeking advice on how to invest their retirement savings outside of their 401k plan.¹⁰ Our auditors were randomly assigned to four different treatment portfolios that reflect different types of investment strategies or biases. We will discuss these strategies in more detail below. We also vary the wealth ranges of the clients, either between US\$ 45,000 and US\$ 55,000 or between US\$ 95,000 and US\$ 105,000. These ranges were picked to mimic the savings for

¹⁰ If the shopper was asked for a 401(k) plan investment, the standardized answer was that a 401(k) plan existed but that she wanted advice on how to invest the extra money. Almost all advisers accepted this statement even though it might lead to an inferior household portfolio. However, if advisers asked for a 401(k) portfolio, they may have first wanted to enlist the client and would later include the 401(k) plan in the portfolio optimization. With respect to owning or renting real estate, we told our auditors to always say that they rented their apartment to avoid situations in which the best advice would be to reduce the mortgage first.

average US households in different age ranges.¹¹ Our study focuses on retail advisers at the lower end of the wealth spectrum, e.g., we do not include private wealth managers or hedge funds. The modal adviser in our study is working either for a bank, retail investment firm, or their own independent operation, focusing on the lower end of the retail segment. Most of them are paid on commission based on the fees and volumes that they generate, and only a small subset of the advisers are independent and would be paid based on capital under management. The fraction of this latter type of adviser is very small in our sample since they usually only deal with wealthier clients.¹²

2.2. Treatments

This setup allows us to test how advisers react to pre-existing investment strategies or client biases. For that purpose, we set up four different treatments to differentiate biases that go against the adviser's self-interest versus those where the adviser's and the client's interests are aligned. We designed the biases in such a way that the net expected loss to the investor is similar in magnitude. As discussed in the introduction, if advisers act purely out of self-interested motives, they should counteract client biases that lead to low-fee income (e.g., excessive investment in company stock) but reinforce biases that increase the adviser's ability to generate fees, such as trend-chasing. However, if advisers are constrained by having to cater to clients' pre-existing beliefs, we should expect that advisers are more restricted in the advice they can give, as in cases, for example, when the client has strong prior beliefs or is emotionally attached to the current portfolio. In contrast, an adviser is less restricted when the client has no predetermined opinion, and thus we should not see a differential response to different types of biases.

To test the importance of these countervailing forces, we selected four different treatments that are presented to the advisers (and impersonated by our auditors). As our "bias scenarios," we selected two of the most common biases documented in the literature: chasing fund returns and investing in employer stocks. We complemented these with two "unbiased scenarios" – a diversified low-fee stock/bond portfolio and an all-cash portfolio.

In *scenario 1* ("chasing fund returns") our auditors indicated that they had been trying to outperform the market by identifying industries that had excess returns in the recent past. In the advice session, the auditor would present the adviser with a portfolio that is concentrated in a few industries with high returns in the last year and ask the adviser to help identify more stocks and industries of this type. Note that de-biasing a client by diversifying the current portfolio would lead to (one-time) returns for the adviser, but he could profit even more by catering to the bias and turning over the portfolio at least once a year. We set up the portfolio such that 30% of the portfolio was invested in one sector exchange traded fund that had performed well in the previous year (i.e. 2007).¹³ These sectors included telecommunication, oil & gas, metals & mining, and US aerospace & defense. Depending on the age group (about 30 or 45 years old), 20% or 35% was invested in intermediate US high-credit quality bond funds. The rest of the portfolio (50% or 35%)

¹¹ In addition, these amounts are varied around the average annual household income in the Boston area (about US\$ 75,000).

¹² Financial advice by independent advisers, who are compensated by the hour or based on capital under management, is often not available or it is too expensive at several hundred dollars for a first visit.

¹³ We fixed the proportion at 30% for two main reasons. First, we wanted to give the adviser the opportunity to invest more in this strategy, although this would lead to less diversification. Second, we believe that the more extreme a portfolio allocation (see also scenario 2 with employer stock), the higher the probability that an adviser would remember a portfolio s/he saw some days ago from another potential client.

was invested in a single S&P 500 index fund. We varied the selection of index funds and the exact amount invested to reduce the probability that an adviser would recognize the portfolio from a previous visit.¹⁴ The average under-performance of the four selected sectors compared to the S&P 500 over 1.5 years *after* the end of our audit study has been about -6.5% p.a., i.e., the 30% investment in this sector resulted in an underperformance between US\$ 1,000 and US\$ 2,000 per year depending on the portfolio size. Even though returns-chasing may be similar to a momentum strategy, we believe that this is not appropriate in the context of this market since investment horizons of customers are rather long and the frequency with which people rebalance their portfolios is too low to allow them effectively to take advantage of momentum strategies.

In *scenario 2* (“employer stock”), we assigned the auditor to one of the 50 largest employers in the Boston area and assigned 30% of the person’s portfolio to company stock.¹⁵ Depending on the age group, 20% or 35% was invested in bonds and the rest in the S&P 500 (as in scenario 1). In this scenario, it was in the interest of both the financial adviser and the client to restructure the existing portfolio. The adviser can earn money in the portfolio rebalancing process, and the client will most likely end up with a better diversified portfolio. To allocate 30% of the portfolio to one stock increases portfolio risk even if we ignore human capital risk. Let’s assume a standard deviation in the market of 20% and a risk premium of 6% p.a., whereas the company stock has a 50% standard deviation, unit beta, and 6% risk premium. Then, a 30% allocation to company stock lowers the Sharpe ratio from 0.3 to 0.247, translating for a given risk to a return loss of 1.29% or US\$ 1,032 per year for an US\$ 80,000 portfolio.¹⁶ Appendix A contains examples of scenario 1 and 2 portfolios and additional client background information. Again, we did not want to assign more than 30% of the portfolio to company stock, since on average people usually do not hold more than 50% of assets in company stock and so that we avoided raising advisers’ suspicion.

In addition to these two scenarios with inherent biases, auditors were assigned a well diversified portfolio in *scenario 3*, consisting of low-fee US index stock funds and bonds, using the same allocation to bonds depending on the age group, as in all other scenarios. While the portfolio is the most efficient of all the treatments used in the study, this treatment does have a (US) home bias, and thus a value-enhancing adjustment to the portfolio might be to suggest more international diversification.¹⁷ Moving the low-fee portfolio to an actively managed portfolio with the same risk/return profile but average management fees would result in additional costs of about one percentage point per year, i.e., between US\$ 500 and US\$ 1,000 in our scenario.

Scenario 4 is our control treatment, since the available money is currently invested in a short-term certificates of deposit and the auditor does not display any preconceived biases. In this scenario, only the investment amount and the demographics are varied as before. The adviser receives no hints how the client would like to invest the money except that the client would like help with a better investment strategy.

¹⁴ Even though we used only ETFs or low-cost index funds both for the diversified part of the portfolio (treatments 1-3)) and for the sectors in treatment 1, only one adviser mentioned this fact.

¹⁵ Auditors were assigned to one of these employers in all treatments but had employer stock in their portfolios only in treatment 2. We provided them with some background information about their employer, including travel times to various points in Boston and Cambridge, such that auditors could answer basic questions about their working life.

¹⁶ We thank John Campbell for this example.

¹⁷ Given the high volatility of currency exchange rates over the last ten years, it is not obvious whether international diversification helps to improve the portfolio characteristics.

Auditors were randomly assigned to one of the four scenarios. They had to be college-educated and had to match our gender and age requirements. One group was in their early 30s, and we assigned them a financial wealth of about US\$ 50,000. The other group is in their mid- to late 40s and had a financial wealth of about US\$ 100,000. All auditors were assigned to one of the 50 largest employers in the Boston area. In scenario 2 (employer stocks), the employer had to be listed on an exchange. All other characteristics (like marital status, children, etc.) that auditors may have talked about with their advisers were their own characteristics, such that they could talk naturally about them. The investment horizon of all auditors was retirement age, i.e., about 30 years for the first group and about 15 years for the second group.

2.3. Logistics of Audit

To implement the actual logistics of the visits, we hired a financial audit firm that specializes in identifying and training auditors. We worked very closely with the firm to select suitable people as auditors for the study, and we were also intimately involved in training the auditors. We designed all the training scripts and set up the schedule of visits with predefined advisers to which auditors were randomized. To ensure that auditors were able to understand the advice that was given to them, they had to know at least some basics of financial products and received some guidelines on how to ask for specific advice. Auditors were trained first about basic financial literacy through our online manuscript. Then, they participated in a training session via video conference with the supervisors and our staff. Finally, audit candidates had to take a short online test to qualify for the study (about 10% of the pre-selected auditors failed and were excluded from this study. See Appendix B). Even after training, it is still possible that auditors do not retain all the information that advisers provide in the meeting. However, such behavior should just lead to more noise overall and not bias our results. Auditors were assigned only to one treatment at a time to avoid confusion and retraining. After the first set of audits was finished, we reassigned auditors to a second treatment to control for auditor fixed effects.

The audit firm provided the logistics of monitoring and implementing the scheduling of visits, setting up online survey forms, and finding and compensating auditors. Auditors were paid on a per visit basis and were told that they would not be invited for a repeat assignment if we heard any complaints about their behavior. We also sent our research associate to do random spot checks in order to observe whether the client was meeting with the adviser. To minimize any demand effects, we made sure that the study was triple blind: the financial advisers, the financial audit firm, and the auditors did not know why and how we chose specific parameters. The company and the auditors were told that the aim of our study was to conduct an assessment of the quality of the market for financial advice and that any variation in the treatment arms was instituted to create variation in order to minimize detection and suspicion.¹⁸

Logistics of the Meeting: To set up the in-person meeting, auditors called their assigned financial adviser and agreed on a time convenient for both.¹⁹ As a reason for the visit, the auditors stated that they were seeking advice on how to invest privately held retirement savings they had outside of employer-provided vehicles (401(k) and defined benefit plans).

¹⁸ Note that we never mentioned different treatments, neither in our conversations with the audit firm nor with the auditors.

¹⁹ A new adviser was only assigned after the previous visit had been completed. Advisers were at most visited once a week by different auditors.

At the agreed-upon time, the auditors would meet with the advisers for a consultation of about one hour, usually in the adviser's office. During the meeting, the auditor would follow the general script provided by us. Depending on their treatment assignment, they would explain their existing investment strategy and ask for advice with their portfolios and investment strategies as described above. The auditors were asked to write down their assigned portfolio on a piece of paper or print them out so that they could show the status quo to the adviser. There was enough variation in the way the information was presented that advisers would not be suspicious of any potential repetitions. Aside from the actual treatment assignment, auditors were told to answer truthfully any information concerning their name, social security number, or any other demographic information, such as number of children or marital status.

Tracking Advice: We encouraged auditors to write down information that the adviser provided in the course of the meeting to increase accuracy. Taking notes is natural in an advice situation and thus does not create any suspicion. Again, we made sure that the auditors had enough variation in how they would put down the notes to avoid any potentially suspicious repetitions. One caveat about the scope of advice is that many advisers are unwilling to provide detailed, personalized advice (e.g., advice on the allocation of assets to specific funds) unless the client has moved his or her funds to the adviser's firm.²⁰ Since our auditors were not able to provide the adviser with those funds, some advisers were reluctant to provide very specific advice and rather commented in general terms about the quality of the clients' existing portfolios and the recommended allocation going forward. Therefore, in most of the study we will focus on the type of advice given and the associated reasoning.

After the visit, auditors were asked to fill out an online exit survey that had multiple choice questions with free-text fields. They had 24 hours to fill out this information after the conclusion of the visit to make sure that they had not forgotten the information they obtained. In addition, each auditor had to send in the adviser's business card such that the audit company could make random calls to verify that the auditor actually had shown up to the visit.²¹ If the questionnaire was not available within 24 hours, the auditor was contacted by a supervisor and reminded to provide the information. This procedure helped extract high-quality and complete information after the visit. Moreover, auditors were only paid after filling in the form. For most questions, a "Don't know/Don't remember" option existed to avoid random answers. If auditors had received additional written information at or after the meeting with their adviser, they forwarded these materials to us and we coded the written recommendations if any were made. There was only one auditor who did not fill out his surveys in the necessary time and was subsequently dropped from the study. We also conducted an exit interview with each auditor after their first visit to verify that auditors were comfortable with the setup.

3. Summary Statistics and Randomization

The audit data of 284 client visits was collected between April and August 2008, i.e. after the problems of Bear Stearns surfaced but before the bankruptcy of Lehman Brothers in mid-September. We had initially planned for an audit of 480 observations but unfortunately had to stop our audit study prematurely, since in the ensuing financial contraction the

²⁰ This behavior is to some extent comparable to a car dealer who asks first for a down payment before agreeing to a test drive of a car.

²¹ Information on the adviser's identity was not passed on to us.

market for financial advice in the Boston area was significantly restructured.²² Moreover, the changing economic conditions were especially important for the *chasing returns* treatment since the outperforming industries of the previous year had changed.

As a result, our four scenarios are not evenly covered. **Table 1** shows the distribution of visits across the four different scenarios: there are 103 visits in scenario 1, 62 visits in scenario 2, 49 visits in scenario 3, and 70 visits in the last scenario. However, **Table 1** confirms that, despite the reduced sample size, the randomization of visits to advisers still seems to be intact. The average age of auditors does not vary across the treatment groups and is centered around 39-40 years. The average assigned annual income is US\$ 80,000 and again there is no significant difference between the four cells. The same is true for the investment amounts; the average investment is between US\$ 77,000 and US\$ 80,000. Finally, on average, the fraction of female auditors is about 77% and there are no significant differences between the different treatment groups.

While the power of the tests is lowered, due to the smaller sample size, it is reassuring that the randomization largely holds despite the smaller sample.

4. Results

4.1. Descriptive Statistics of a Typical Audit

Table 2 provides an overview over the information the adviser collected and her recommendation during the auditor's visit. It is a prerequisite to understand the client's financial situation, their ability to handle portfolio risk, and their current exposure to market risk through their other investments. Some of the basic information that an adviser should ask about are the income level of the client, whether they have savings in a 401(k) plan apart from the money they want to invest with the adviser, their occupation, and whether they have children. We form indicator variables equal to one if the adviser asked for the specific information at some point in the consultation and zero otherwise. The results in the first five rows of **Table 2** show that in the vast majority of cases advisers start off by asking the auditor for basic personal characteristics such as age, income, whether they have children, and whether they have a 401(k) plan.²³ On average these questions are asked in more than 70% of the visits.²⁴

In **Table 3**, we regress a dummy indicating whether the adviser asked about the client's age on the gender and log age of the client. We focus on these two characteristics since they are most easily observable from the outset of the visit. We find that women were asked for their age less often while the coefficient on log age is not significant but positive. Similarly, in columns (2) through (4), we see that women auditors were asked about their personal and financial situation less often than men. Columns (3) and (4) of **Table 3** show that older people were asked about their financial situation concerning income and whether they have a 401(k) plan more often than younger people. These results could indicate that advisers adjust their approach towards a potential client to reflect their expectation about the future return from this client including the probability of recruiting her: older auditors have higher investment funds, and men are usually viewed as being more willing to move their account to

²² Financial advice firms started to consolidate their advisory business by reducing the number of advisers. Thus, arranging visits within our design, given the previous visits, became almost impossible.

²³ As mentioned above, auditors said that they are investing in a 401(k) plan but do not want to discuss the details of these investments.

²⁴ Advisers ask for this crucial information even though it is not a legal requirement as in other countries.

another adviser. In both cases, the expected return for the adviser is higher than with younger or female clients. As a consequence, advisers are less likely to ask younger or female auditors some basic question about their financial situation, and it also leads to worse advice since the adviser does not have full information.

When we look at the advisers' recommendations in the same table, we see that advisers had a much higher propensity to suggest actively managed mutual funds than index funds. The advisers encouraged the client to invest in index funds in only 7.5% of the advice sessions (21 visits). In contrast, in 50% (or 142) of the visits, the adviser suggested investing in actively managed funds. This stark discrepancy is a first indication that advisers might be trying to guide clients to high-fee investments.²⁵ In that context, it is interesting to see that a majority of advisers mention (some) fees of the recommended funds spontaneously, without the client having to ask for it. But in many cases, the fee discussion is used to downplay the impact of fees. Given that clients usually come to advisers to receive help with their investments, it might not be too surprising that on average advisers try to change the clients' fund allocation. What is interesting, however, is that they tended to move shoppers away from the existing strategy regardless of the initial portfolio, i.e., even when they looked at a low-fee diversified portfolio. So they were willing to make the client effectively worse off.

Another interesting finding, inferred from clients' free-form answers, is that some advisers (84 visits, or roughly 30%) refused to offer any specific advice until the auditor transferred resources to the adviser. This is interesting because it illustrates a screening problem for customers: it is hard for them to judge the value of an adviser before committing to the adviser.

Finally, the mix of asset allocations is interesting (see **Table 6**). Advisers recommended on average an investment of about 63.5% in equity, 23.8% in bonds, and 12.7% in cash. However, the mean advice suggests an international equity allocation of about 27% of the portfolio. While this is a smaller international allocation than optimal portfolio theory suggests, this is an aggressive equity allocation and in that sense leans against any bias that might generate an equity premium.

4.2 Advisers' reactions to the current investment strategy

In a second step, we examine how treatment assignments affect the advice that auditors received from the advisers. As discussed above, since we randomly assign clients and their portfolios to different advisers, we can test the average response of a typical adviser without any concern of clients' self-selection to different types of advisers. First, in **Table 4** we focus on the adviser's overall reaction to the auditor's assigned portfolio as a function of the different scenarios presented by the auditors. We measure whether the adviser is supportive of the auditor's current strategy and supports more investment in the same portfolio or is critical of the current strategy and discourages further investment in the same.²⁶ This analysis excludes the "cash only" treatment. In that treatment, it is not possible

²⁵ Note that catering to client's beliefs based on their current portfolio would lead advisers to recommend more and/or different index funds since 70% (treatment 1 and 2) or 100% (treatment 3) of the current portfolio are invested in index funds.

²⁶ The encouragement/support variable is based on the question "Did the adviser make any comments about how you should modify your existing portfolio?" Auditors could choose from three multiple choice answers:

for the adviser to express support for the “strategy” since the client explicitly asks for help in improving the investment strategy. If advisers try to de-bias their clients, they should not be supportive of the trend-chasing or employer stock strategy but supportive of the diversified portfolio strategy. At most, they should suggest some international diversification, in that case. In contrast, if advisers aim to maximize their fee income they should be supportive of trend-chasing but not supportive of the employer stock or efficient portfolio strategy, since these will not generate a high number of transactions and fees.

Each cell in **Table 4** presents the mean, the standard deviation, and the p-value of a t-test between the subset presented in that cell versus the sample mean. Overall, advisers seem to support strategies that result in more transactions and higher management fees. The first two rows show the mean responses for whether the adviser supported the current investment strategy or suggested to change it.²⁷ In the case of the returns-chasing portfolio, the advisers were significantly more supportive than for either the company stock or the index portfolio. The likelihood of a supportive response was 19.4% for the returns-chasing portfolio, against the sample mean of 13.1%, but only 9.7% for the company stock portfolio and a remarkably low 2.4% in the case of the index portfolio. When we measure whether the adviser proactively encouraged the client to change the current investment strategy, we see a parallel pattern. The incident of a negative response is significantly below the mean for the returns-chasing portfolio but significantly above the mean for the index fund portfolio: in 59.2% of cases, an adviser suggested a change in the current strategy. In contrast, if the client had an index portfolio, the adviser suggested changing the current investment strategy in 85.4% of the cases.

These results show that even though the meeting between the client and adviser is also a sales situation, advisers are willing to go against the (revealed) preferences of the client and suggest changes away from the current strategy in the majority of cases. Therefore, the catering concerns are less strong in this context than one might have conjectured. However, it is possible that the visits start off differently from the ultimately suggested strategy and reflect more of the sales pitch: in the next two rows, we analyze the adviser’s *initial* response to the auditor’s portfolio as opposed to the final recommendations we saw before. This first reaction could be interpreted as a judgment of the client’s prior behavior. We indeed see in row (3) that the adviser’s initial reaction to the different portfolios is most likely to be positive in the case of the index portfolio (24.4%), least positive for the company stock portfolio (12.9%), and about average for the returns-chasing portfolio (16.5%). Explicitly negative comments about the portfolio are observed only in 23.3% of the cases with returns-chasing portfolio but in about 56% for both the index and company stock portfolios.

These results show advisers try to be more positive initially, even though they ultimately want to change the clients’ strategies. Interestingly, advisers seem more compelled to say something explicitly complementing about the index portfolio and are less comfortable to do so for the returns-chasing or the company stock portfolios. It is possible that advisers

(a) “Adviser encouraged me to invest more in the existing strategy”,

(b) “Adviser said that I should not change the allocation in my existing strategy but not invest more” or

(c) “Adviser discouraged me to invest more in the existing strategy”.

The encouragement variable is coded as 1 if the auditor answered (a) and 0 otherwise. Negative comments (discouragement) to the same question are coded as 1 if the auditor chose (c) and 0 otherwise. The correlation between the two dummy variables is -0.5374.

²⁷ We have too many banks and too few repeat visits in our sample such that we cannot include bank fixed effects. Probit regressions or using random effects instead of clustering at the auditor level have qualitatively the same results and are available on request.

feel constraint by what they know is the better advice, or that they do not want to undermine their credibility by being outright dismissive of diversification and index funds, assuming that most people have at least heard through the media and other sources that these are good investment guidelines.

Table 5 contains the results of a parallel analysis but using a regression set-up, where we cluster observations at the individual client level to control for base-rate characteristics. The regressions control for the age and income level of the auditors, since these were the two variables on which we stratified auditors. We also include month fixed effects since the study was implemented over a five months period to reduce the likelihood of detection. Column (1) reports the results from a regression of a dummy indicating encouragement on the treatment dummies where the omitted category is the returns-chasing portfolio (scenario 1). The results show that advisers were least supportive of the efficient portfolio, followed by the company stock treatment, with coefficients of $-.28$ and $-.17$, respectively, i.e., advisers supported the returns-chasing strategy significantly more often than the other two strategies.²⁸

In column (2), we replicate the regression setup of column (1) but include additional characteristics of the clients such as gender, marital status, and investment amount. None of the characteristics are significant and including these controls does not change the coefficients on the treatment dummies, as is to be expected given random assignment. In regressions (3) and (4), we now break out customers with an assigned investment amount of about US\$ 100,000 and about US\$ 50,000. The idea is to see if the advice across the scenarios differs for wealthier clients, and we find no support for any such differences.

In columns (5) to (8), we repeat the regression setup of the first four columns but use as a left-hand side variable whether the adviser strongly discouraged investing further in the existing strategy. We again find consistently with the prior results that advisers were most negative on the efficient portfolio. The coefficient is $.4$ and significant at the 1% level. There is no significant difference for company stock, however. These results strongly suggest that advisers try to dissuade clients from investing in an efficient portfolio, likely because this minimizes the fee income for the adviser. Interestingly, this incentive seems to be so strong that advisers are willing to push clients out of investments portfolios that are close to perfectly efficient (index scenario). However, it does seem that the investor is more constrained in the case of company stock, perhaps because of an attempt to cater to the clients' bias. This produces a perverse situation where the adviser is actively leaning against an efficient portfolio but not willing to lean against what is actually a biased strategy.

In columns (9) and (10), we look at the initial reactions of the adviser. The difference between the dependent variable here and in columns (1) through (4) is that we now focus on the adviser's very first reaction to the client's portfolio, i.e., looking backward. In contrast, the dependent variable in columns (1) through (4) focuses on how the adviser argued the portfolio should be structured or restructured going forward. As expected based on our results in **Table 4**, we now find very different advisers' reactions. Initially, advisers react significantly more positively to the index portfolio relative to the returns-chasing portfolio. But there is no significant difference between company stock and returns - chasing. In addition, these differences are much less pronounced than in the first eight columns. All this suggests that advisers are more moderated in their initial reactions. And

²⁸ In these regressions we use scenario 2 as the benchmark case to detect differences between scenarios 2 and 3. These results are available on request.

they seem much less willing to make overtly negative comments about the clients' prior choices.

4.3 Asset Allocation

In the next step, we evaluate the overall asset allocation (stocks, bonds, domestic or international investments) that is recommended by the advisers. In a large fraction of the visits, we did not get very detailed quantitative advice about which specific funds to invest in, since many advisers insisted that the client should first place the funds with the adviser's company. Therefore, we can compare differences in the capital allocation across broad asset classes. In the analysis in **Table 6**, we will now include all four different scenarios including the cash treatment (scenario 4), since we look at the recommendations that the adviser makes for the portfolio going forward. The mean comparison in **Table 6** is parallel to the setup in **Table 4**. The outcome variables here are measured as the recommended percentage of the portfolio that should be invested in one of the respective asset classes. The first three rows report the results for the fraction of assets in stock or bonds (we omit cash as the residual category) and the fraction of assets to be invested internationally. As can be seen in row 4, we lose about 45% of the sample in this analysis since many advisers did not provide specific enough recommendations.²⁹ Second, the means tests reported in **Table 6** indicate that advisers did not seem to tailor the asset allocation according to the scenario that auditors come in with, since none of the comparisons are even close to being significant. The one exception is that auditors who came in with an index portfolio seem to have received a higher suggested allocation to bonds on average across the sample advisers, suggest investing 23.8% of the assets in bonds, but this fraction goes to 31.9% for clients arriving with an index portfolio. This is surprising since our control group might be seen as being even more risk averse since they have had no exposure to equities. While we cannot rule out that this might be a spurious correlation, it could possibly suggest that advisers assessed people who hold index portfolios as more risk averse and thus felt that a stronger allocation to bonds could be in line with the auditor's preferences. But at the same time this interpretation is not so easy to square with the fact that advisers were strongly pushing the clients who come in with index portfolios to move out of these and into actively managed funds.³⁰

Advisers did not mention fees at a significantly different frequency depending on our four treatments (row 5). However, there are significant differences between treatments with respect to recommending actively managed or index funds (rows 6 and 7). The two most concrete dimensions of advice that we can measure are whether the adviser recommended actively managed funds and/or (passive) index funds as an investment to the client. These two dimensions are of interest since a large body of literature on mutual fund returns suggests that actively managed funds on average have lower net returns but allow fund companies and advisers to charge higher fees. In contrast, index funds have been shown to be a better investment option for retail investors since they provide access to

²⁹ We also used an indicator variable for whether the adviser did at all mention investment in any of the above mentioned asset classes and the results are qualitatively very similar.

³⁰ Another explanation is related to the amount of fees that can be earned with stock funds compared to bond funds. As a consequence, it does not pay off to know more about bond funds than about stock funds, i.e. advisers simply do not know enough to recommend another bond funds. Advisers may also believe that it is easier to move clients from index bond funds into actively managed stock funds than shifting them to actively managed bond funds.

investing in a well diversified portfolio at a low fee structure (see for example Gruber 1996). Rows 6 and 7 of **Table 6** reconfirm that on average advisers were much more likely to recommend actively managed funds (49.7% of the cases) versus index funds (only 7.3% of the cases). Interestingly, however, the likelihoods of when advisers suggest one of these two types of funds varies with the scenario: actively managed funds were recommended especially frequently to auditors who came in with an index fund portfolio (61.0% of the time)³¹ or just cash investments (75.0% of the time). In the case of returns-chasing portfolios, actively managed funds were suggested only in 40.8% of the visits and even less frequently for company stock portfolios, where it was only 24.2% of the time. In contrast, index funds were almost never mentioned by the advisers. Only when people came in with a cash portfolio the adviser was significantly more likely to recommend index funds.

As before, we now use a regression framework to confirm the robustness of the results. Column (1) of **Table 7** follows the usual baseline setup where we regress the recommended fraction of bonds on the dummies for the four scenarios (again, scenario 1 is the omitted one) and controls for the client characteristics. The fraction of bonds that was recommended for the client's portfolio does not seem to vary with the different scenarios. The same is true in columns (2) through (4) when looking at the allocations to the other asset classes. Again, the advice does not seem to vary by scenario. The one exception is that clients in the all-cash scenario 4 have significantly lower exposure to stocks (column 2). This could suggest that clients who are in scenario 4 are considered to be very risk averse (or even unsophisticated) by the advisers and therefore the advisers might think that these clients would not be able to handle the risk exposure of a high fraction of equities. But the results are quite noisy, which is most likely due to the much smaller sample size.

The recommended investment in stocks and domestic assets significantly increased with clients' annual income, which may be explained by an assumed higher risk or loss tolerance. However, the recommended exposure to stocks decreased with larger investment amounts. Married clients were advised to have significantly more bond and stock investments at the expense of liquidity, whereas female clients were advised to invest significantly less in both asset classes based. The months following the collapse of Bear Stearns had an effect on recommended domestic investments only – clients were told in May through July to invest domestically significantly less.

These different recommendations based on personal characteristics may be caused by the adviser's information collection process. Note that clients always disclosed the investment amount at the beginning of the conversation since they asked for advice with respect to their current portfolio. The likelihood of being asked for their current occupation (regression (2)) or annual income (regression (3)) significantly decreased with higher investment amounts (see **Table 3** in section 4.1).

4.4 Investment suggestions: Adviser and client incentives

We now also run the above results with respect to advisers' recommendations of actively managed or index funds in regression specification to make sure that they survive a number of different controls. In column (1) of **Table 8**, we regress an indicator variable equal to one if the adviser recommended investing in index funds and zero otherwise on dummies for

³¹ This result may not be too surprising at first. But recommending an internationally well-diversified low fee index funds would have been an option for the adviser and would have counted as "recommending an index fund".

the four scenarios with the returns-chasing scenario as the omitted category and controls for the age and income level of the client. All the regressions are clustered at the auditor level. The results in column (1) show that advisers were significantly more likely to recommend index funds to clients who came in with an “all cash” portfolio relative to the returns-chasing treatment (the coefficient on the “all cash” scenario is 0.18 and is significant at the 5% level). The coefficients on all other scenarios are close to zero and statistically insignificant. In column (2), we repeat the same regression but include further controls for client characteristics such as gender, marital status, number of children, and the amount they client wants to invest. The results are unchanged compared to column (1). In columns (3) and (4), we break out the sample into the visits where auditors were assigned an investment amount around US\$ 100,000 versus US\$ 50,000.³²

In contrast, in columns (5) to (8) of **Table 8**, we repeat the same regression setup as before but the dependent variable now is a dummy for whether the adviser suggested actively managed funds to the clients. It is important to note that advisers recommended actively managed funds in about half of all visits while they only recommended index funds in about 8% of visits. The baseline regression in column (5) shows that advisers were much more likely to recommend actively managed funds to clients who came in with either the index fund portfolio (scenario 3) or the all-cash portfolio (scenario 4). Instead, advisers were almost 20% less likely to mention actively managed funds to clients in the company stock portfolio (scenario 2). Again, the results are unchanged with additional controls for client characteristics in column (6). As before, we then break out the sub-sample of clients that have about US\$ 100,000 to invest (column 7) and those that have about US\$ 50,000 to invest (column 8). Interestingly, clients with higher investment amounts tend to be recommended actively managed funds in scenarios 3 and 4 at a much higher rate than clients that are less wealthy. Moreover, less wealthy clients received significantly fewer recommendations in the employer stock scenario, which may reflect advisers’ beliefs that the probability of de-biasing these clients is rather low.

Overall, advisers seemed to maximize fees by placing more weight on actively managed funds that created a higher fee income. Most strikingly, even if a client had a well-diversified index funds portfolio, the adviser encouraged investment in actively managed funds. The objective of the adviser in this behavior might have been to signal that they could add value to the client by suggesting something different from the existing portfolio. This behavior was particularly pronounced for wealthier clients where the fee income mattered more to the adviser. But advisers could also have achieved this goal by suggesting low-fee international diversification. In general, advisers did not proactively reach out to clients to rebalance the portfolio due to changing circumstances of the client, but only to sell them new funds and generate fees. The advice that we observe in our treatments are a good proxy for the different situations that an adviser might encounter with their clients throughout a longer term relationship. The evidence suggests that most of the interaction is driven by the need to generate fees rather than to respond to the clients rebalancing needs.

But advisers also seem to attempt to cater to their clients’ perceived preferences. For clients that came in with a company stock portfolio, advisers were much less likely to suggest actively managed funds. One could hypothesize that high concentration in company stock might suggest to the adviser that this client is more risk averse and passive in their investment approach and thus might not be comfortable investing in an actively managed

³² The percentage improves to 10.7% (from 4.4%) in the index funds scenario when we eliminate those visits in which advisers are not willing to provide any advice before the client transfers her portfolio.

fund. Given that actively managed funds were recommended to half of the clients, it is no surprise that advisers spontaneously mentioned fees just as often or as seldom in all scenarios and with respect to most control variables. However, advisers spontaneously addressed fees with an increasing likelihood if clients were older and with a smaller investment amount (regression (10)). The first result can be explained by an assumed higher experience level, i.e., advisers pre-empted an often heard question in this age group. The second result may seem surprising, given that less wealthy clients received more recommendations for actively managed funds. But mentioning fees can involve talking about load fees (or discount of load fees) as well as management fees, which are more important in the long run. Thus, advisers may be more likely to talk about the less relevant fees with lower levels of wealth and less likely to talk about (all) fees with higher levels of wealth.

4.5 Personal Characteristics and Customization of Advice

In **Table 9**, we first analyze the advice that was given as a function of client characteristics. Characteristics such as gender and age can be inferred relatively easily by meeting the client. We also include characteristics of the client's personal situation that the adviser learns from the conversation, such as the client's marital status, the number of children, and income level. Standard finance theory would suggest that advisers should take personal circumstances of the client into account when setting the investment strategy. For example, clients with a shorter investment horizon, such as older clients, most textbooks would suggest that they should not be investing in risky long-term strategies. Similarly, risk averse clients should not be guided towards risky investments, such as having a high percentage in of their portfolio of stocks or actively managed funds.

In column (1), we investigate whether the adviser encouraged the auditor to invest in index funds as a function of the auditor's characteristics. Standard errors in all regressions are clustered at the auditor level. We see that the coefficient on log of age is positive but not significant. Similarly none of the other coefficients on the client characteristics are significant. Thus, there are no discernible differences between clients that were encouraged to invest in index funds. But it is important to keep in mind that the overall incidence of suggesting index funds was very low (8%). In column (2), we repeat the same regression set-up but use an indicator variable for whether the adviser suggested actively managed funds as the dependent variable. Now, the coefficient on log of age is large and statistically significant. Older clients were more often encouraged to invest in actively managed funds. Similarly, people with children were encouraged to invest in actively managed funds. In contrast, the coefficient on the gender dummy is negative and highly significant. This suggests that advisers are less likely to recommend actively managed funds to female clients. This could be in line with the belief that women are more risk averse and thus would have a lesser tolerance for actively managed funds. Alternatively, women are perceived as being more fee-sensitive.

We find that the only significant and economically very large coefficient of whether the adviser spontaneously mentioned the fees is log age (see column (3)). Advisers tend to explain the fee structure of funds much more to older clients than to younger ones. None of the other coefficients on characteristics are significant. The belief must be that older people are more cost-conscious and potentially better informed about investment options. Therefore, advisers proactively discuss the fees rather than trying to ignore the topic.

Finally, in column (4), we analyze the adviser's attitude towards the client relationship. We learned through the free text answers of our auditors that some advisers refused to offer any specific advice as long as the potential client had not transferred his or her account to the company of the adviser (n=84). The intention of the adviser seems to be that they first wanted to sign up the client before parting with any useful information. On the one hand, it makes sense that advisers would want to protect their time and insights so that clients do not replicate the advice for free. However, this situation is puzzling, since it forces the client to choose an adviser without being able to get any indication of the person's quality upfront. The result in column (4) shows that this behavior was most pronounced towards female auditors. Advisers were almost 40% more likely to tell female clients that they first had to transfer the balance to them. One could imagine that this behavior might be based on the perception that women are more docile or gullible.

5. Conclusions

Overall our findings suggest that the market for financial advice does not de-bias retail investors and if anything may exaggerate existing biases. While advisors seem to take into account client characteristics such as age or family status when making portfolio recommendations, they are unwilling to lean against biases that help them further their own economic interest, e.g. maximize fees. We find that in some cases the advice even pushes clients towards funds with higher expected fees with little change in portfolio diversification and thus would reduce the expected returns on their portfolios (as discussed in the case of the index portfolio). Intermediation in this market on average does not seem to correct individual biases. Therefore, these biases can have first order implications for aggregate capital flows and possibly pricing of risk more broadly, if there is not enough informed capital to exploit arbitrage opportunities against capital flows from "biased" retail investors. While in times of normal economic activity these biases might be arbitrated away, one could worry that in times of crisis they become particularly important when arbitrage capital dries up. It can also shed light on how we should model information aggregation in equilibrium if competitive forces in the market for financial advice do not lead to the provision of the best possible advice. Competition might be limited by the fact that financial advisers exploit the biases of naïve (or uninformed) retail investors. At the same time, advisers who are interested in providing better advice might be unable to gain market share if biased retail investors are unable to differentiate advice that is in the interest of the client from advice that is only in the self-interest of the advisor. While we cannot rule out that catering to client beliefs plays some role, our evidence suggests that adviser self-interest plays an important role in generating advice that is not in the best interest of the clients. Our findings show that advisers do not feel constraint to only give investment recommendations that affirm the clients' prior beliefs for fear of losing the client. In fact, advisors have no problem discouraging clients from investing more in their current strategies if this is not in the interest of the advisor. But we do see some evidence that the advisers are fully aware that this is a sales situation, since in the initial interaction with the client they always praise the client's prior investment choices no matter what they invested in.

These results are intriguing but they are also only a first step in what is a very important research area. They open the door to a set of other questions about the market for advice that our current sample size or treatments do not allow us to answer. Three questions stand out. First, does the nature and quality of the advice depend on the adviser's incentives? Answering this question will require a larger sample of advisers with varying

incentives, e.g., fee-only advisers. Though such advisers have grown in recent years, they are still a minority in the market. More importantly, auditing them would raise the costs of audits significantly since each audit would now entail a fee. While our current tests reflect the advice that the representative adviser in the market provides, it would be interesting to understand how variation in incentives might affect the interaction with the clients. For example, it is possible that fee-only advisers provide better advice but have to charge such high fees that average retail investors are better off in the status quo situation. Second, does the quality of advice change over time? In an ongoing and evolving advice relationship, an adviser might either use the relationship to tailor the advice more closely to the clients' needs or to be tempted to take advantage of the increasing trust. Third, how does the demand for financial advice look like and what role does it play in shaping the type of advisers that survive in the market? Specifically, how do individuals assess the quality of advice? In this audit study, auditors were willing to go back to about 70% of the advisers they visited but now with their own money (see **Table 2**). In other words, most advisers succeeded in convincing their potential clients and thus they have no need to change their advice giving. To understand why clients like the currently available financial advice would require a different methodology where the unit of observation would need to be potential investors. These types of questions are essential going forward to understanding the forces that shape the equilibrium in the market for advice.

References

- Altonji, Joseph G. and Rebecca M. Blank, 1999, Race and gender in the labor market, In Orley C. Ashenfelter and David E. Card (eds.), *Handbook of Labor Economics*, edition 1, volume 3, chapter 48, 3143–3259, Amsterdam: Elsevier Science.
- Barber, Brad M. and Terrance Odean, 2000, Trading Is Hazardous To Your Wealth: The Common Stock Investment Performance of Individual Investors, *Journal of Finance* 55, 773–806.
- Barberis, Nicholas and Richard H. Thaler, 2003, A Survey of Behavioral Finance. In G. Constantinides, M. Harris, and R. Stulz (eds.), *Handbook of Economics of Finance*, 1052–1121, Amsterdam: Elsevier Science.
- Benartzi, Shlomo and Richard H. Thaler, 2001, Naive Diversification Strategies in Retirement Saving Plans, *American Economic Review*, 91.1 (March), 79–98.
- Benartzi, Shlomo and Richard H. Thaler, 2007, Heuristics and Biases in Retirement Savings Behavior, *Journal of Economic Perspectives* 21.3 (Summer), 81–04.
- Bergstresser, Daniel, John M.R. Chalmers, and Peter Tufano, 2009, Assessing the Costs and Benefits of Brokers: A Preliminary Analysis of the Mutual Fund Industry, *Review of Financial Studies* 22, 4129–4156.
- Bhattacharya, Utpal, Andreas Hackethal, Simon Kaesler, Benjamin Loos, and Steffen Meyer, 2012, Is Unbiased Financial Advice To Retail Investors Sufficient? Answers from a Large Field Study, *Review of Financial Studies* (forthcoming).
- Bluethgen, Ralph, Andreas Gintschel, Armin Müller, and Andreas Hackethal, 2008, Financial Advice and Individual Investors' Portfolios, Goethe Universität Frankfurt am Main.
- Bodie, Zvi and Dwight B. Crane, 1997, Personal Investing: Advice, Theory, and Evidence from a Survey of TIAA-CREF Participants, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=36158.
- Bogle, John C., 2000, *John Bogle on Investing: The First 50 Years*, New York: McGraw-Hill.
- Bolton, Patrick, Xavier Freixas, and Joel Shapiro, 2007. Conflicts of Interest, Information Provision, and Competition in the Financial Services Industry, *Journal of Financial Economics* 85.2 (August), 297-330.
- Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of finance*, Vol. 52, No. 1, pp. 57-82.
- Campbell, John Y., 2006, Household Finance, *Journal of Finance* 61, 1553–1604.
- Canner, Niko, N. Gregory Mankiw, and David N. Weil, 1997, An Asset Allocation Puzzle, *American Economic Review* 87.1 (March), 181–191.
- Chalmers, John and Jonathan Reuter, 2011, What is the Impact of Financial Advice on Retirement Portfolio Choice and Outcomes?, Working paper.
- Chater, Nick, Roman Inderst, and Steffen Huck, 2010, Consumer decision-making in retail investment services. Report to the European Commission (SANCO).
- Chen, Hailiang, De, Prabuddha, Hu, Yu Jeffrey and Hwang, Byoung-Hyoun, 2011, Customers as Advisors: The Role of Social Media in Financial Markets, doi:10.2139/ssrn.1807265.
- Del Guercio, Diane, Jonathan Reuter and Paula Tkac, 2010, Demand for Financial Advice, Broker Incentives, and Mutual Fund Market Segmentation, NBER Working Paper No. 16312.
- Fix, Michael and Margery A. Turner, 1998, *A National Report Card on Discrimination in America: The Role of Testing*, Washington, DC: Urban Institute Press.
- Georgarakos, Dimitris and Roman Inderst, 2011, Financial Advice and Stock Market Participation,

ECB Working Paper No. 1296.

- Gruber, Martin, 1996, Another Puzzle: The Growth in Actively Managed Mutual Funds, *Journal of Finance* 51.3 (July), 783-810.
- Hackethal, Andreas, Roman Inderst, and Steffen Meyer, 2011, Trading on Advice, doi:10.2139/ssrn.1701777.
- Hackethal, Andreas, Michael Haliassos, and Tullio Jappelli, 2012, Financial Advisors: A Case of Babysitters?, *Journal of Banking and Finance*, 36(2), 509-24.
- Heckman, James J., 1998, Detecting Discrimination, *Journal of Economic Perspectives* 12.2 (Spring), 101-16.
- Hung, Angela A. , Noreen Clancy, Jeff Dominitz, Eric Talley, Claude Berrebi, and Farrukh Suvankulov, 2008, Investor and Industry Perspectives on Investment Advisers and Broker-Dealers, Technical Report, RAND Center for Corporate Ethics and Governance.
- Inderst, Roman and Marco Ottaviani, 2009, Misselling through Agents, *American Economic Review* 99.3 (June), 883-908.
- Inderst, Roman and Marco Ottaviani, 2011a, Competition through Commissions and Kickbacks, mimeo.
- Inderst, Roman and Marco Ottaviani, 2011b, How (not) to pay for advice: A framework for consumer financial protection, mimeo.
- Iyer, Rajkamal and Antoinette Schoar, 2009, The Role of Culture in Financial Negotiations: Evidence from an Audit Study, *International Differences in Entrepreneurship*, NBER Conference Volume (forthcoming).
- Iyer, Rajkamal and Antoinette Schoar, 2011, Ex Post (In)Efficient Negotiation and Breakdown of Trade, mimeo.
- Kramer, Marc, 2012, Financial Advice and Individual Investor Portfolio Performance, *Financial Management* (forthcoming),
- Levitt, Steven D. and Chad Syverson, 2008, Market Distortions when Agents are Better Informed: The Value of Information in Real Estate Transactions, *Review of Economics and Statistics* 90.4 (November), 599-611.
- Malmendier, Ulrike and Devin Shanthikumar, 2007, Are small investors naive about incentives?, *Journal of Financial Economics* 85.2 (August), 457-489.
- Moore, Don A., George Loewenstein, Daylian Cain, and Max H. Bazerman (eds.), 2005, *Conflicts of Interest*. Cambridge University Press.
- Schneider, Henry S., 2009, Agency Problems and Reputation in Expert Services: Evidence from Auto Repair, mimeo.

Table 1: Randomization across Scenarios

Scenarios	Chasing Returns PF (1)	Company Stock PF (2)	Index Funds PF (3)	All Cash (4)
Number of Observations	103	62	49	70
Age	39	40	40	40
Annual Income	US\$81,000	US\$80,000	US\$82,000	US\$81,000
Investment Amount	US\$77,000	US\$81,000	US\$80,000	US\$79,000
% Female Auditors	77%	80%	78%	75%

This table shows the distribution over the four different scenarios of the 284 audit visits to financial advisers between April and August 2008. In scenario 1, 30% of the client's portfolio (PF) was invested in a sector that outperformed the S&P 500 in the year prior to the audit study (Chasing Returns). In scenario 2, clients' current portfolios contained 30% company stock. Scenario 3 refers to a diversified client portfolio with only low-fee index funds. These funds were also used for the other 70% of the portfolio in scenario 1 and scenario 2. In scenario 4, the client said that all the money was invested in a certificate of deposit at a local bank. The auditor's actual age was used whereas the annual income and the investment amount were assigned by the study.

Table 2: Descriptive Statistics about the Adviser-Client Conversation

VARIABLES	Yes	No	Total
Advisers encourage more of current strategy?	28	186	214
Advisers suggest change of current strategy?	144	70	214
Initial Reaction Positive	36	178	214
Initial Reaction Negative	85	129	214
Recommend index funds	21	263	284
Recommend actively managed funds	142	142	284
Spontaneously mention fees	160	121	281
Ask about age	236	48	284
Ask about current occupation	217	67	284
Ask about annual income	212	72	284
Ask about 401k	252	32	284
Ask about number of children	200	84	284
Auditor would go back to this adviser with own money	200	84	284

This table contains descriptive statistics for the 284 audit visits at financial advisers. Auditors in the control treatment did not answer the first four questions. As answers to the question "Did the adviser make any comments about how you should modify your existing portfolio?" auditors could choose from multiple choice answers:(a) "Adviser encouraged me to invest more in the existing strategy", (b)"Adviser said that I should not change the allocation in my existing strategy but not invest more" or (c)"Adviser discouraged me to invest more in the existing strategy". The answer is counted as an encouragement if the auditor answered (a). If the adviser picked answer (c), he suggested a change of the current strategy. An initial positive or negative reaction to the clients' current portfolio has been recorded, too. Auditors also entered in the online exit questionnaire whether index funds or actively managed funds were recommended and whether any fees were mentioned by the adviser without being asked. In addition, auditors recorded whether they were asked at some point during their visit about their age, current occupation, annual income, the existence of a 401(k) plan or children. Finally, we asked auditors after each visit whether they would go back to this adviser with their own money, or not.

Table 3: Personal characteristics and adviser's information gathering

VARIABLES	(1)	(2)	(3)	(4)
	her age	her current occupation	her annual income	a 401(k) plan
log(Auditor's Age)	0.127 (0.153)	0.205 (0.169)	0.342** (0.165)	0.321** (0.141)
Gender	-0.171* (0.101)	-0.106 (0.0656)	-0.182* (0.0931)	-0.107** (0.0525)
Constant	-5.234* (2.737)	-7.865*** (2.785)	-5.709** (2.201)	-1.243 (1.765)
Observations	283	283	283	283
R-squared	0.178	0.129	0.111	0.127

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

This table shows regression results for advisers' questions about the auditor's age, her current occupation, her annual income and the existence of a 401(k) plan (all variables yes: 1/no: 0) based on 284 audit visits at financial advisers between April (omitted) and August 2008. The additional explanatory variables are the auditor's actual age her gender (female=1) as dummy variables.

Table 4: Advisers' initial reaction and (non-) supportive recommendations in different treatments

	Overall Sample	Chasing Returns	Company Stock	Index Funds
Is the adviser <i>supportive</i> of the current strategy?	13.11	19.41 (39.74) 0.007	9.68 (29.81) 0.0341	2.44 (15.62) 0.024
Is the adviser <i>against</i> the current strategy?	67.48	59.22 (49.38) 0.011	69.35 (46.48) 0.707	85.37 (35.78) 0.006
Initial reaction positive	16.99	16.50 (37.30) 0.854	12.90 (33.79) 0.308	24.39 (43.48) 0.160
Initial reaction negative	39.81	23.30 (42.48) 0.001	56.45 (49.98) 0.001	56.09 (50.24) 0.017
#observations	214	103	62	49

This table contains data collected in 214 visits (without the control treatment 4) at a financial adviser. Reported are the means, standard deviations and p-values of tests between the sample mean and the treatment mean. The two variables (supportive/not supportive) are coded based on the question "Did the adviser make any comments about how you should modify your existing portfolio?" Auditors could choose from multiple choice answers: (a) "Adviser encouraged me to invest more in the existing strategy", (b) "Adviser said that I should not change the allocation in my existing strategy but not invest more" or (c) "Adviser discouraged me to invest more in the existing strategy". The answer is counted as an encouragement (=1) if the auditor answered (a) and 0 otherwise. The answer is seen as discouraging (=1) if the auditor chose (c) and 0 otherwise. An initial positive (yes: 1/no: 0) or negative (yes: 1/no: 0) reaction to the clients' current portfolio are used for rows (3) and (4), respectively.

Table 5: Advisers' initial reaction and (non-) supportive recommendations

VARIABLES	(1) adviser encouraged client to invest more in the existing strategy				(2) adviser discouraged client to invest more in the existing strategy				(9) positive initial reaction	(10) negative initial reaction
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)		
Investment Amount	≈\$100k		≈\$50k		≈\$100k		≈\$50k			
Company Stock PF	-0.165** (0.0807)	-0.185** (0.0780)	-0.216 (0.154)	-0.205* (0.102)	0.110 (0.0908)	0.162 (0.111)	0.0550 (0.179)	0.297** (0.117)	0.0137 (0.0766)	0.252* (0.125)
Index Funds PF	-0.284** (0.111)	-0.304** (0.116)	-0.418* (0.210)	-0.300** (0.108)	0.397*** (0.0989)	0.438*** (0.109)	0.476** (0.176)	0.521*** (0.147)	0.199** (0.0884)	0.195 (0.167)
log(Auditor's Age)	0.0965 (0.152)	0.0121 (0.396)	0.991 (0.732)	-0.681 (0.489)	-0.0175 (0.181)	0.380 (0.485)	0.394 (0.481)	1.888** (0.775)	0.327* (0.181)	0.265 (0.275)
log(Annual Income)	-0.266 (0.228)	-0.249 (0.260)	-0.124 (0.349)	-0.801*** (0.166)	0.358 (0.252)	0.381 (0.262)	0.123 (0.390)	0.751** (0.275)	0.194 (0.357)	-0.0901 (0.353)
log(Investment Amount)		0.0190 (0.258)				-0.186 (0.293)				
Marital Status		-0.0354 (0.0355)				0.0866 (0.0924)				
Children		-0.125* (0.0682)				0.135 (0.119)				
Gender		-0.0119 (0.108)				-0.0871 (0.127)				
Constant	2.612 (2.273)	2.561 (2.185)	-2.026 (3.556)	11.24*** (2.984)	-3.911 (2.781)	-3.539 (3.039)	-2.442 (3.308)	-14.86*** (4.765)	-3.287 (4.006)	0.0809 (3.494)
Observations	214	214	105	109	214	214	105	109	214	214
R-squared	0.093	0.116	0.137	0.158	0.107	0.127	0.148	0.165	0.075	0.172

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

This table shows regression results for advisers' reactions to the clients' current portfolios based on 284 audit visits at financial advisers between April (omitted) and August 2008 -- did they encourage the client to pursue the current strategy or did they discourage the client? The two variables are coded based on the question "Did the adviser make any comments about how you should modify your existing portfolio?" Auditors could choose from multiple choice answers (a) "Adviser encouraged me to invest more in the existing strategy", (b) "Adviser said that I should not change the allocation in my existing strategy but not invest more" or (c) "Adviser discouraged me to invest more in the existing strategy". The answer is counted as an encouragement (=1) if the auditor answered (a) and 0 otherwise. The answer is discouraging (=1) if the auditor chose (c) and 0 otherwise. An initial positive (yes: 1/no: 0) or negative (yes: 1/no: 0) reaction to the clients' current portfolio are used as dependent variables for regressions (9) and (10). The explanatory variables are dummy variables for the first three scenarios: In scenario 1 (omitted), clients were chasing fund returns. In scenario 2, clients' current portfolios (PF) contained 30% company stock. Scenario 3 refers to a diversified client portfolio with only low-fee index funds. The auditor's actual age, her annual income assigned by the study, and the investment amount assigned by the study are used, as well. Finally, the client's actual marital status (married=1), the existence of at least one child in the household, and the client's gender (female=1) enter as dummy variables.

Table 6: Average advisers' asset allocation recommendations and fees

	Overall Sample	Chasing Returns	Company Stock	Index Funds	All Cash
allocation bonds %	23.79 (16.16)	23.71 (13.07)	21.84 (15.64)	31.94 (19.42)	21.65 (16.43)
		0.964	0.369	0.007	0.258
allocation stock %	63.48 (25.97)	67.13 (22.85)	63.65 (26.21)	67.27 (19.91)	57.89 (30.59)
		0.220	0.964	0.411	0.056
allocation international %	26.81 (17.21)	29.46 (14.91)	24.67 (19.49)	23.23 (11.06)	27.78 (20.45)
		0.199	0.414	0.234	0.662
observations (allocation)					
adviser mentions fees spontaneously	56.94 (49.61)	59.41 (49.35)	48.39 (50.38)	60.53 (49.53)	58.75 (49.53)
		0.533	0.124	0.638	0.701
actively managed funds	49.65 (50.09)	40.78 (49.38)	24.19 (43.18)	60.98 (49.38)	75.00 (43.57)
		0.024	0.000	0.117	0.000
index funds	7.34 (26.13)	2.91 (16.89)	4.84 (21.63)	0.00 (0.00)	18.75 (39.27)
		0.031	0.395	0.051	0.000
observations (all)	284	103	62	49	70

This table contains data collected in 284 visits at financial advisers in the Boston/Cambridge area in 2008. Reported are the means, standard deviations and p-values of tests between the sample mean and the treatment mean. There are fewer observations of the allocation data since not all advisers provided allocation advice. Mentioning fees (yes: 1/no: 0), actively managed funds (yes: 1/no: 0) and index funds (yes: 1/no: 0) are used for rows 5, 6 and 7, respectively.

Table 7: Advisers' asset allocation recommendations

VARIABLES	(1)	(2)	(3)	(4)
	adviser recommends to invest % of portfolio			
	in bonds	in stocks	internationally	domestically
Company Stock PF	-0.351 (4.569)	0.0278 (5.735)	-4.347 (6.797)	-0.306 (7.785)
Index Funds PF	7.108 (4.963)	3.776 (6.792)	-3.595 (6.094)	10.24 (6.699)
All Cash PF	-2.435 (6.046)	-9.851* (5.143)	-3.595 (8.691)	-13.06 (8.925)
log(Auditor's Age)	7.866 (8.620)	-0.812 (12.66)	5.149 (10.34)	-2.739 (14.09)
log(Annual Income)	12.57 (11.96)	27.52** (12.72)	-19.63 (15.05)	33.42* (19.23)
log(Investment Amount)	-2.444 (1.526)	-3.982*** (0.905)	-0.127 (0.766)	-2.437* (1.293)
Marital Status	0.908 (3.890)	20.44*** (3.804)	10.92** (4.676)	14.26* (7.364)
Children	-6.962 (4.221)	6.187 (4.496)	4.131 (6.536)	-15.49* (7.871)
Gender	-11.57** (4.314)	-9.621* (5.252)	-11.31** (4.970)	-10.76 (8.484)
Constant	-116.7 (150.3)	-205.4 (156.4)	237.6 (177.0)	-251.3 (255.5)
Observations	167	172	152	152
R-squared	0.174	0.211	0.128	0.334

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

This table shows regression results for advisers' portfolio allocation recommendations for bonds, stocks, internationally and domestically (all in %) based on 284 audit visits at financial advisers between April (omitted) and August 2008. The explanatory variables are dummy variables for all four scenarios: In scenario 1 (omitted), clients were chasing fund returns. In scenario 2, clients' current portfolios (PF) contained 30% company stock. Scenario 3 refers to a diversified client portfolio with only low-fee index funds. In scenario 4, the client claimed that all the money was invested in a certificate of deposit at a local bank. The auditor's actual age, her annual income assigned by the study, investment amount assigned by the study. In addition, we use the client's actual marital status (married=1), the existence of at least one child in the household, and the client's gender (female=1) as dummy variables.

Table 8: Advisers' recommendations and mentioning fees

VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	adviser recommends index funds				adviser recommends actively managed funds				adviser talks spontaneously about (any) fees			
Investment Amount			≈\$100k	≈\$50k			≈\$100k	≈\$50k			≈\$100k	≈\$50k
Company Stock PF	0.0166 (0.0540)	0.0157 (0.0567)	0.0222 (0.0246)	0.102 (0.179)	-0.198** (0.0906)	-0.149* (0.0737)	-0.00673 (0.0796)	-0.455** (0.180)	-0.108 (0.0966)	-0.0967 (0.103)	-0.0133 (0.152)	-0.0774 (0.157)
Index Funds PF	-0.0191 (0.0354)	-0.0218 (0.0415)	0.0442 (0.0429)	0.105 (0.100)	0.255* (0.138)	0.308** (0.119)	0.663*** (0.0978)	-0.101 (0.167)	-0.0149 (0.142)	-0.0444 (0.133)	0.372* (0.215)	-0.248 (0.146)
All Cash PF	0.183** (0.0827)	0.168** (0.0697)	0.256*** (0.0658)	0.181 (0.137)	0.318*** (0.104)	0.292*** (0.105)	0.517*** (0.0704)	0.0750 (0.144)	-0.0106 (0.114)	-0.0274 (0.107)	0.0860 (0.167)	0.0421 (0.118)
log(Auditor's Age)	-0.0847 (0.120)	0.129 (0.110)	-0.541 (0.410)	1.559 (0.969)	0.287 (0.172)	0.504** (0.193)	-0.231 (0.560)	1.600 (1.050)	0.325 (0.196)	0.494** (0.205)	-0.563 (0.911)	1.190 (0.782)
log(Annual Income)	0.0313 (0.141)	0.0292 (0.126)	0.142 (0.119)	-0.165 (0.183)	-0.170 (0.291)	-0.267 (0.273)	-0.0744 (0.272)	0.186 (0.458)	-0.134 (0.240)	-0.131 (0.220)	0.223 (0.401)	-0.526 (0.312)
log(Investment Amount)		-0.109*** (0.0110)				-0.0783*** (0.0220)				-0.0620*** (0.0175)		
Marital Status		0.0271 (0.0365)				0.00512 (0.0623)				0.106 (0.0889)		
Children		-0.0521 (0.0518)				0.272*** (0.0913)				-0.0638 (0.0997)		
Gender		0.0553 (0.0490)				-0.134* (0.0744)				0.0409 (0.101)		
Constant	-0.0528 (1.349)	0.321 (1.285)	0.217 (1.847)	-3.570 (3.211)	0.884 (3.311)	2.188 (2.968)	2.067 (3.105)	-7.605* (4.085)	0.849 (2.862)	0.934 (2.673)	0.0365 (5.195)	2.371 (4.307)
Observations	284	284	148	136	284	284	148	136	281	281	148	133
R-squared	0.101	0.177	0.185	0.144	0.180	0.226	0.301	0.211	0.028	0.049	0.070	0.094

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

This table shows regression results for advisers' investment recommendations of index funds or actively managed funds (yes: 1/no: 0) based on 284 audit visits at financial advisors between April (omitted) and August 2008. In addition, regression results are shown for advisers' mentioning of fees (yes: 1/no: 0) without being prompted by the client. The explanatory variables are dummy variables for all four scenarios: In scenario 1 (omitted), clients were chasing fund returns. In scenario 2, clients' current portfolios (PF) contained 30% company stock. Scenario 3 refers to a diversified client portfolio with only low-fee index funds. In scenario 4, the client claimed that all the money was invested in a certificate of deposit at a local bank. The auditor's actual age, her annual income assigned by the study, investment amount assigned by the study. In addition, we use the client's actual marital status (married=1), the existence of at least one child in the household, and the client's gender (female=1) as dummy variables.

Table 9: Advice as a function of auditor characteristics

VARIABLES	(1) adviser recommends index funds	(2) adviser actively manages funds	(3) adviser mentions fees	(4) adviser offers advice only after money transfer
log(Auditor's Age)	0.190 (0.118)	0.643*** (0.217)	0.527** (0.197)	-1.283 (0.811)
log(Annual Income)	-0.0683 (0.120)	-0.361 (0.272)	-0.106 (0.219)	0.488 (0.350)
Marital Status	0.0557 (0.0366)	0.107 (0.0876)	0.130 (0.0804)	-0.124 (0.105)
Children	-0.0203 (0.0438)	0.320*** (0.0813)	-0.0471 (0.101)	0.138 (0.151)
Gender	0.0194 (0.0417)	-0.209** (0.0979)	0.0146 (0.101)	0.383*** (0.116)
Constant	1.366 (1.248)	2.884 (3.068)	0.574 (2.688)	-11.20** (4.488)
Observations	284	284	280	128
R-squared	0.113	0.124	0.045	0.218

Robust standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

This table shows regression results for advisers' recommendations of index funds and actively managed funds based on all 284 audit visits at financial advisers between April (omitted) and August 2008. The other regressions evaluate when advisers spontaneously mention (any) fees and why advisers offer advice only after the portfolio or money is transferred to the new account. The additional explanatory variables are the auditor's actual age, her assigned annual income, her actual marital status (married=1) and the client's gender (female=1) as dummy variables.

Appendices – Audit study design examples and test questions for potential auditors

Appendix A – Portfolio and background information examples

Scenario 1 (chasing fund returns)

Shopper ID	3		
Scenario	1	Company	Haemonetics
Gender	Female	Occupation	Product Development
Age	30-35	Annual Income	\$65.000
		\$ to Invest	\$45.000
		Portfolio ID	1T1
		Portfolio	\$45.000,00
		SPDR S&P Metals & Mining (XME)	\$13.500,00
		Vanguard 500 (VFINX)	\$22.500,00
		Vanguard Interm-Term Bond Index (VBIIIX)	\$9.000,00
			On-line account on E-Trade (mention only if asked)

Scenario 2 (employer stocks)

Shopper ID	52		
Scenario	2	Company	Analog Devices
Gender	Male	Occupation	Intellectual Property
Age	30-35	Annual Income	\$75.000
		\$ to Invest	\$49.000
		Portfolio ID	1T2
		Portfolio	\$49.000,00
		Analog Devices Stocks (ADI)*	\$14.700,00
		Vanguard 500 (VFINX)	\$24.500,00
		Vanguard Interm-Term Bond Index (VBIIIX)	\$9.800,00
			On-line account on E-Trade (mention only if asked)

Appendix B – Test for potential auditors

1. Please give an example of a fund family?

- a. General Mills
- b. Vanguard
- c. US Government
- d. Sovereign Bank

2. What do advisers mean when they talk of actively managed funds?

- a. Funds where the manager picks securities in which to invest the money
- b. Any fund that has a manager
- c. Funds that track the market portfolio

3. Usually index funds have lower fees than actively managed funds?

True/false

4. The definition of dividends is

- a. the profits that a company retains each year to invest in new projects
- b. a name for the increase in annual stock price of a company
- c. the part of profit that a company pays to its investors each year

5. Which of the following is not a “security”?

- a. Procter and Gamble stock
- b. Dividend from Coca-Cola
- c. US Government bond
- d. Oppenheimer Capital Appreciation A (OPTFX)

6. Advisers will usually recommend that your portfolio should be diversified between bonds, stocks and mutual funds. Imagine you had US\$ 20,000 to invest and wanted to invest 60% in equity and 40% in bonds, which of the following portfolios should you pick?

- a. US\$ 6,000 in Vanguard S&P 500 index
US\$ 4,000 in MFS high Income Municipal
US\$ 10,000 in Certificate of Deposit
- b. US\$ 8,000 US Government bonds
US\$ 1,000 Ford Motor stock
US\$ 6,000 in Vanguard S&P 500 index
US\$ 5,000 in Alliance Bernstein Intl Growth A
- c. US\$ 8,000 MFS high Income Municipal
US\$ 10,000 IBM stock
- d. US\$ 8,000 Vanguard 500 Index
US\$ 12,000 US Government Bonds

7. What is the relation between risk of default and return of a corporate bond?

- a. The higher the risk of default the lower should be the interest rate
- b. The lower the risk of default the lower the interest rate
- c. More trustworthy debtors offer higher interest rates
- d. Return is independent of risk

8. A front load is

- a. a sales commission the investor has to pay periodically for holding a fund
- b. a fund with a heavy concentration in short term securities
- c. a measure of the operational costs of the fund
- d. a sales charge the investor pays when buying a mutual fund

75



**THE ECONOMIC IMPLICATIONS OF THE DEPARTMENT OF LABOR'S
2010 PROPOSALS FOR BROKER-DEALERS**

Alicia H. Munnell, Anthony Webb, and Francis M. Vitagliano

CRR WP 2013-4
Date Released: February 2013

Center for Retirement Research at Boston College
Hovey House
140 Commonwealth Avenue
Chestnut Hill, MA 02467
Tel: 617-552-1762
<http://crr.bc.edu>

All of the authors are affiliated with the Center for Retirement Research at Boston College (CRR). Alicia H. Munnell is the director of the CRR and the Peter F. Drucker Professor of Management Sciences at Boston College's Carroll School of Management. Anthony Webb is a senior research economist and Francis Vitagliano is a consultant for the CRR. The research reported herein was pursuant to a grant from the Ford Foundation. The findings and conclusions expressed are solely those of the authors and do not represent the views of the Ford Foundation or Boston College. The authors would like to thank Andrew Eschtruth, Tray Spilker, and Zhenyu Li for their substantial contributions to the project.

© 2013, Alicia H. Munnell, Anthony Webb, and Francis M. Vitagliano. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

About the Center for Retirement Research

The *Center for Retirement Research at Boston College*, part of a consortium that includes parallel centers at the University of Michigan and the National Bureau of Economic Research, was established in 1998 through a grant from the Social Security Administration. The Center's mission is to produce first-class research and forge a strong link between the academic community and decision-makers in the public and private sectors around an issue of critical importance to the nation's future. To achieve this mission, the Center sponsors a wide variety of research projects, transmits new findings to a broad audience, trains new scholars, and broadens access to valuable data sources.

Center for Retirement Research at Boston College
Hovey House
140 Commonwealth Avenue
Chestnut Hill, MA 02467
phone: 617-552-1762 fax: 617-552-0191
e-mail: crr@bc.edu
crr.bc.edu

Affiliated Institutions:
The Brookings Institution
Massachusetts Institute of Technology
Syracuse University
Urban Institute

Abstract

In 2010, the U.S. Department of Labor (DOL) proposed changes that would eliminate third-party incentive payments, such as 12b-1 fees, that may encourage broker-dealers to sell high-fee mutual funds to Individual Retirement Account (IRA) customers. The investment industry argues that eliminating these fees could force broker-dealers to charge directly for advice, which could result in less advice being provided and customers making poor investment decisions. This paper examines the tradeoff between lower fees and investor mistakes due to forgoing advice. The analysis draws on existing pricing and cost evidence from the United States and on the experience of ongoing reforms in the United Kingdom aimed at lowering fees and improving investment advice. An educated guess suggests that the elimination of 12b-1 fees would reduce IRA customer costs by 4 basis points. If broker-dealers responded by moving customers away from high-cost, actively-managed funds into low-cost index funds, IRA customers could save another 7 basis points on their IRAs. Additional savings might be possible to the extent that actively-managed funds underperform their relevant indices. It is unlikely that broker-dealers would change their business model with respect to the provision of advice as a result of the loss of 12b-1 fees. The paper also uses an inter-temporal optimization model to quantify the potential benefits and costs of reform in the United States. The results of the optimization exercise point to relatively modest potential benefits, but even more modest costs under plausible assumptions. Given the size of the policy problem, several more extensive reforms should be considered to ensure that retirement saving in tax-advantaged accounts is invested more effectively. These ideas include: making it easier to retain accumulated assets in the 401(k) system; 2) making the rollover from a 401(k) to an IRA an ERISA-covered event; 3) extending ERISA to all rollover IRAs; and 4) instituting changes to further control fees in both 401(k)s and rollover IRAs.

Introduction

Both the U.S. Department of Labor (DOL) and the U.S. Securities and Exchange Commission (SEC) are considering changes that would affect the conduct of broker-dealers serving retail clients. (The DOL issued a formal regulatory proposal in 2010 and is expected to issue a new proposal shortly; the SEC staff made recommendations to the Commission in a 2011 report.) The motivation for the DOL changes is to reduce the likelihood that third-party incentive payments encourage broker-dealers to sell high-fee mutual funds that will substantially reduce ultimate asset accumulations for customers with Individual Retirement Accounts (IRA). The SEC has the broader concern that investors do not recognize the differences in standards to which investment professionals are held and wants to level the playing field.

The two agencies, which administer different statutes, necessarily take different approaches. The SEC would extend fiduciary conduct standards to broker-dealers who provide investment advice to both retirement and non-retirement accounts. The SEC change would mean that broker-dealers must act “in the best interest” of the customers, as opposed to the current standard of “suitability.” The new standard would be somewhat higher. The SEC might also require broker-dealers to disclose whether they have a potential conflict of interest – for example, whether they will receive a commission from the provider for selling a particular product.

The 2010 DOL proposal, by broadening the activities that would be considered the provision of advice, would extend *different* fiduciary obligations to anyone who gives advice to IRAs (banks, insurance companies, Registered Investment Advisers, and broker-dealers).¹ This paper focuses on broker-dealers because they account for the bulk of IRA investments and their compensation arrangements involve the most prohibited transactions. The DOL proposal would not require a change in the standard of conduct toward clients, which makes it different from the SEC proposal. Rather, it would make more broker-dealers “fiduciaries” under the Internal Revenue Code (IRC) and thus subject to IRS prohibited transaction rules. Importantly, under the self-

¹ The DOL proposal would also extend fiduciary obligations to broker-dealers advising 401(k) participants through self-directed brokerage accounts. But given that only a tiny fraction of 401(k) accounts have a self-directed component, this part of the proposal is not the major source of controversy.

dealing provision, it would eliminate third-party fees, such as 12b-1 and other revenue-sharing fees. The 12b-1 fees are paid by mutual fund providers to broker-dealers for marketing, distribution, and servicing expenses; they continue as long as the customer holds the shares; and they provide a clear incentive for broker-dealers to sell high-fee products.

The 2010 DOL proposal, which has been met with a storm of criticism from the investment industry, is the primary focus of this paper. The thrust of the criticism is that a large number of IRAs, especially those with smaller balances, are brokerage accounts. Brokers handling these accounts offer advice in the process of performing transactions, and that advice is often paid for by third-party payments such as 12b-1 fees. According to the industry, eliminating these fees could force brokers to charge directly for their advice by instituting an explicit charge in the form of a percent of assets under management. The industry asserts that such a pricing approach would raise costs generally and, by increasing the visibility of the cost of advice, result in less advice being provided for low- and moderate-income IRA holders.

Thus, the impact of the proposed reform involves assessing the trade-off between: 1) the beneficial effects of lowering fees by eliminating 12b-1 fees and reducing the incentive to sell high-cost, actively-managed funds; and 2) the deleterious effects on saving rates and portfolio allocation of any reduction in the availability of advice.

The paper proceeds as follows. The first section describes the nature of the problem – namely, that, by rolling over money from 401(k)s to IRAs, more people are coming into contact with broker-dealers, and broker-sold products are often expensive. The second section discusses the current regulatory environment and the proposed changes. The third and fourth sections assess the likely impact of the proposed DOL reform on fees, on the provision of advice, and on the industry. The third section makes an educated guess about the potential impact of the proposed DOL reform, drawing on the experience to date of the United Kingdom's Retail Distribution Review, where the evidence suggests that the reform – far more radical than that suggested by the DOL – has lowered fees but not substantially reduced the availability of advice. The fourth section uses an inter-temporal optimization model to identify the benefits and possible costs of reform. The results point to relatively modest potential benefits from eliminating 12b-1 fees, but

also show that costs will outweigh benefits only under quite implausible assumptions as to the availability of advice and the mistakes made by those forgoing advice. In short, the modest DOL proposal will reduce fees slightly, have virtually no effect on the business model of broker-dealers, and any diminution in access to advice, which is likely to be small, will have little effect.

Given the modest impact of the DOL proposal, the fifth section offers several more extensive reform options – some could be accomplished through rule-making and some would require legislation. The most extensive would be to subject roll-over IRAs to ERISA’s fiduciary standards for those rendering advice, as the money in these accounts comes from the employer plan arena. The same legislation could ban actively-managed mutual funds – and their high fees – from both 401(k) plans and the newly classified IRAs. A less ambitious proposal would clarify that the transition from a 401(k) plan to an IRA is an ERISA-covered event and set a low-fee target date fund as the default within the IRA. Other steps could be taken to keep the money in the 401(k) system, such as requiring employers to accept rollovers from any former 401(k) plan.

The overall conclusion is that additional protections are required in the IRA market. As long as accumulations are held in 401(k) plans, participants are operating in a world in which sponsors must operate as fiduciaries and fees are under a spotlight. Once they roll over their accounts into IRAs, they enter a world where suitability becomes the standard of care and broker-dealers are paid commissions that encourage the sale of high-priced mutual funds. If a fiduciary standard and attention to fees are appropriate for retirement assets when they are in the plan, such safeguards are clearly still appropriate when they are rolled over. The DOL recognizes this logic, but, with its authority limited to defining who is a fiduciary under the IRC, has put forth a very modest proposal.

The Nature of the Problem

The problem is that the average individual who rolls over his 401(k) plan into an IRA enters a world in which broker-dealers face incentives to sell high-fee investments. Fees have a significant effect on how much an individual will have at retirement: an additional 100 basis

points over a 40-year period reduces final assets by about one fifth.² The problem needs to be addressed because the number of people rolling over into IRAs has increased dramatically, the distinctions between the players in the retail market have become blurred, and the evidence suggests that broker-sold assets are expensive.

Increase in IRA Customers

The demand for IRAs has grown significantly in the wake of the shift in retirement plans from defined benefit to defined contribution – typically 401(k)s. In 1983, the majority of those with an employer-sponsored plan were covered by a defined benefit plan; by 2010, the vast majority relied solely on a 401(k) (see Figure 1). While the changing nature of plans has been widely recognized, the recent shift from 401(k)s to IRAs has received less attention. The increase in IRAs has occurred because individuals roll over their balances when they shift jobs during their worklives and when they withdraw their funds at retirement. By 2010, according to the Federal Reserve's *Survey of Consumer Finances*, 44 million individuals had an IRA (see Table 1). And total IRA assets now exceed the money in 401(k)s (see Figure 2).

The rollover of balances from 401(k)s to IRAs is extraordinary given that participants are typically passive in their interactions with their 401(k) plans. They rarely change their contribution rate or rebalance their portfolios in response to market fluctuations or as they age.³ Because inertia has been identified as a major problem, reforms have focused on automatic mechanisms, such as automatic enrollment, automatic increases in the default contribution rate, and target date funds.⁴ Thus, one would think that the force of inertia would lead participants to leave their balances in their 401(k) accounts until they draw them down in retirement. The fact that they actually take the trouble to move their funds suggests a strong motivating force. Some households may be attracted by the opportunity to obtain a wider menu of investment options or to consolidate their account holdings. But others may be seduced by advertisements from

² The calculations assume real stock and bond returns of 7 percent and 3 percent respectively, a stock asset allocation of two thirds, 40 years of savings, and real wage growth of 1.1 percent per year. If individuals respond to the decline in projected balances by saving more, the ultimate impact on wealth at retirement will be smaller.

³ Munnell and Sundén (2004); and Ameriks and Zeldes (2001).

⁴ Madrian and Shea (2001); Thaler and Benartzi (2004); and Young (2012).

financial service firms urging participants to move their funds out of their “old,” “tired” 401(k) plan into a new IRA.⁵

The assumption by participants must be that the firms advertising rollovers are operating in the participants’ interest, but, in fact, participants are very often moving from being protected by a fiduciary, low-fee environment into a relatively unprotected and potentially high-fee arena.⁶

Blurring of Distinction between Retail Financial Service Providers

The landscape for retail financial services used to be quite simple. Investment advisers provided expert assistance for retail customers in selecting financial instruments.⁷ A separate group of traditional broker-dealers facilitated the buying and selling of financial instruments and any advice they provided was incidental. Investment advisers were required to register with the SEC under the Investment Advisers Act or with their state securities authority, while traditional broker-dealers were not.⁸ The retail customers were typically higher-income individuals, who had their retirement provided by defined benefit plans and were investing supplementary resources.

At the same time that the number and type of people dealing with broker-dealers has increased, the distinction between investment advisers and broker-dealers has become blurred. Banks, mutual funds, stockbrokers, investment advisers, insurance brokers, and financial planners offer

⁵ A Charles Schwab ad shows a man with a 1980s boombox and the tag line “Let’s talk about that 401(k) that you picked up back in the ‘80s.” Merrill Edge (launched by Bank of America, owner of Merrill Lynch) depicts a woman with her arms spread and the phrase “Catching up with my old 401(k)s.” TD Ameritrade shows a sad young woman with writing in the background that says “roll over your old 401(k).” Fidelity’s “follow the green line” campaign includes an ad with a woman speaking to a Fidelity representative about how to roll over her “old 401(k).”

⁶ This situation generally holds for 401(k) participants at large firms. 401(k) participants at small firms sometimes face higher fees, in which case they may be able to reduce expenses by rolling over 401(k) assets to an IRA.

⁷ Frankel (2011) argues that financial markets have grown considerably more complex and that the “contract model” that has prevailed for 60 years should therefore be eliminated, and broker-dealers should, instead, be held to a fiduciary standard.

⁸ The SEC website sets out the criteria for state registration: “Investment advisers that are prohibited from registering with the Commission (e.g. advisers that do not have assets under management of \$25 million) generally must register with the state(s) in which they transact advisory business (e.g., have advisory clients or have a place of business), unless they are exempt from investment adviser regulation under state law. These advisers will be regulated primarily under state law administered by state securities authorities, rather than federal law administered by the SEC.”

a wide variety of investment products and services. In some cases, the roles performed by these service providers overlap. For example, a financial planner who trades securities for a client must also be a broker-dealer. Brokers no longer merely execute trades, but also advise on asset allocation and portfolio selection. Many brokers through advertising and by referring to themselves as “financial planners” or “financial consultants” suggest that their services include planning or consulting services that also involve expert advice. Many individuals and firms that call themselves “financial advisers” may not be considered investment advisers under the law; instead they are regulated as broker-dealers.⁹

Since the IRA market is the retail segment that falls under the DOL’s interpretive authority, it is useful to look at the types of arrangements that individuals have with their financial service provider. Basically, IRA holders investing in stocks, bonds, and mutual funds can have an advisory or a brokerage relationship or both.¹⁰

Advisory IRAs include those accounts with ongoing services, such as investment-specific advice, portfolio allocation recommendations, active monitoring, and investment suitability assessments. These accounts involve a relationship with an investment adviser or a dual-registered broker, generally are fee-based, and tend to represent the most expensive model. Data from a survey by industry consultants suggest that the share of IRAs using an advisory relationship is miniscule for those with low balances and increases to 34 percent for those with balances greater than \$250,000 (see Figure 3).

Brokerage accounts come in two forms: full-service and discount. Full-service account holders have access to a registered broker on an infrequent, point-of-transaction basis. Brokers receive payment in the form of *direct* transaction-based commissions, periodic account fees, and *indirect* commissions through mutual fund loads and 12b-1 fees.¹¹ These payments are typically less than those in advisory accounts but more than those charged by discount brokers.

⁹ Bromberg and Cackley (2012).

¹⁰ As noted in the introduction, a minority of IRA assets (18 percent as of 2012:1 according to the Federal Reserve’s Flow of Funds) are invested with depository institutions, credit unions, and insurers.

¹¹ 12b-1 fees, discussed further below, are paid to brokers by mutual funds from management fees levied on investors.

The discount brokerage model offers the least in financial advisory services and access. Investment decisions are typically self-directed with access limited to branch and call centers and a library of tools to help the investor transact online. Fees within the discount brokerage model are the lowest of the three and are incurred typically through transaction-specific commissions, modest periodic account fees, and indirect charges.

Each component of the IRA market will be affected differently by the DOL's proposal to limit third-party fees. Individuals dealing with registered investment advisers, who are regulated by the SEC, and those investing through discount brokers, whose fees are already low, will be the least affected.¹² It is the middle group – those who have full-service brokerage accounts – that will be most impacted. As discussed below, they could gain, for example, from a reduction in fees as 12b-1 fees are rebated or eliminated and as they are directed to low-cost index funds. These gains, however, could be offset if making the cost of advice more transparent leads to a decline in advice and poor decisions by investors concerning asset allocation (how a portfolio is allocated between asset classes) and portfolio selection (which securities are held within each asset class). Unfortunately, the industry data do not indicate how IRAs are distributed between discount and full-service brokers.

To the extent that the concern is the impact of the change on low- and middle-income households, it is useful to look at who holds IRAs by household income. Roughly one third of households with IRAs fall in the bottom three quintiles of the income distribution (see Table 2). These lower income households are likely to have smaller balances and a brokerage rather than an advisory relationship.

High-Cost Investments

When investing, households need to know: 1) how much to save; and 2) how to invest their

¹² While the statement is generally true, some Registered Investment Advisers receive compensation that would be prohibited if they newly became fiduciaries under a new DOL rule and some of the computer model advice given through discount brokers could become fiduciary advice.

savings to optimize the trade-off between risk and return. This trade-off involves making decisions about asset allocation and portfolio selection. Although financial advisers can perform a valuable service by advising households how much to save, how to diversify, and how to select an appropriate asset allocation, in the current system this advice appears to come at high cost.

Mullainathan, Noeth, and Schoar (2012) report that instead of steering clients towards low-cost index funds, advisers often encourage investment in high-cost, actively-managed funds. Many studies have also shown that actively-managed funds underperform index funds, even before accounting for the higher fees charged by the former.¹³ But broker-sold mutual funds perform worst of all. Bergstresser, Chalmers, and Tufano (2009) estimate that broker-sold funds underperform average actively-managed stock funds by 23 to 255 basis points a year.

Cumulated over a 40-year working career, the disparity between the performances of broker- and direct-sold funds substantially reduces wealth accumulated by retirement. Further, Gil-Bazo and Ruiz-Verdu (2008, 2009) argue that fee-setting by the brokerage community involves charging higher fees on funds with lower expected returns to extract surpluses from unsophisticated investors. Researchers found that brokerage-advised accounts for participants in the Oregon University System's defined contribution plan underperformed self-directed accounts and target date funds while taking on more risk. The underperformance was significant – 154 basis points – and the total difference is even larger once fees are taken into account (Chalmers and Reuter, 2012).

Used as stock and mutual fund pickers, brokers and financial advisers can destroy value. But it does not necessarily follow that households would be better off on their own. Bergstresser, Chalmers, and Tufano (2009) hypothesize that brokers may perform valuable services, not only advising households on saving rates and asset allocation, but also putting a brake on behavioral biases, for example by dissuading clients from panic-selling during market downturns. It is difficult to test these hypotheses because households using brokers may differ in terms of financial knowledge, risk aversion, and temperament from those who invest on their own. But a later section of this paper presents an optimization model that calculates how large the impact on

¹³ For example, see Malkiel (1995, 2005).

these behavioral biases would need to be in order to offset the impact of fees.

Thus, as noted in the introduction, the impact of the proposed reform involves a trade-off between: 1) lower costs as a result of eliminating 12b-1 fees and getting rid of incentives to invest in high-cost, actively-managed funds; and 2) any harm created by a reduction in the availability of advice on saving rates and asset allocation.¹⁴

The Current Regulatory Environment and Proposed Changes

For households holding IRAs, the providers of financial services may well look very similar, but they are regulated by two different entities that apply different standards of conduct.

The Current Regulatory Environment

Investment advisers are regulated by the SEC under the Investment Advisers Act of 1940, and the courts have construed the legislation to require that they act as fiduciaries when providing advice to their clients. The fiduciary standard, which applies to the adviser's entire relationship with his clients, imposes a duty of loyalty and care.¹⁵ The duty of loyalty requires that the fiduciary act in the best interest of the client regardless of the financial implications for the fiduciary. The duty of care requires that fiduciaries investigate to ensure they are providing accurate and complete information to their clients. An adviser who has a material conflict of interest must either eliminate that conflict or fully disclose all material facts relating to that conflict. Investment advisers have traditionally derived most of their revenue from higher

¹⁴ Industry advocates have argued that, under the DOL's proposal, investors would end up paying "twice" for investment advice. If the investor's advisor cannot be compensated via 12b-1 fees from the fund he sells or revenue sharing from the fund's adviser, then he will have to charge the investor directly for his advice. But the fund will not lower its expense ratio, so the investor will still pay the same 12b-1 fees and management fees. (See, for example, Fischel and Kendall 2011). This argument is unpersuasive for two reasons. First, it ignores the likelihood that the advisor who charges directly will simply recommend a lower cost alternative fund. Second, while it may be true that one advisor shifting from indirect compensation to direct charging would have no impact on funds' 12b-1 or management fees, outlawing such indirect compensation for a large market segment should create competition that would push such fees down.

¹⁵ The fiduciary standard applies to the entire advisory relationship, but dual registrants may act as both an adviser and a broker for the same client, operating under different rules when providing different services in connection with different accounts.

income/higher net worth clients and are typically compensated on the basis of assets under management.

Registered representatives of broker-dealers are regulated under the Securities Exchange Act of 1934 through a self-regulatory organization, the Financial Industry Regulatory Authority (FINRA). Brokers and dealers historically were excluded from the definition of advisers under the Investment Advisers Act and thereby from having to comply with the Act's fiduciary standards so long as: 1) the provision of advice is 'solely incidental' to the conduct of business as a broker-dealer; and 2) no 'special compensation' is received for the advisory services. Over the years, however, broker-dealers began to drift into the domain of activities of investment advisers; financial planners emerged as a group of providers that offered a wide range of services; and discount brokers and fee-based accounts further muddied the waters.¹⁶ In response to the blurring of the lines, the SEC in 1999 issued a notice of proposed rulemaking and in 2005 finalized rules that clarified the circumstances under which investment advice from a broker-dealer would be 'solely incidental.' Soon thereafter, the Financial Planning Association challenged the new rule in court and prevailed.¹⁷ The SEC rescinded the rule and began a review of the regulation of broker-dealers and investment advisers.

Broker-dealers must meet a standard of suitability when providing information about financial products and are not required to act solely in the interest of their customers, as specified under ERISA for fiduciaries of employer-sponsored plans. But they do have an obligation to treat their customers fairly, consistent with standards of their profession: their recommendations must be reasonable given their customer's financial situation; they must provide timely and accurate information; and they must disclose conflicts of interest, and so forth.¹⁸ The standard of suitability for broker-dealers raises the possibility that they or their representatives can recommend products that are not necessarily in the interest of the client, but might be considered potentially suitable given the customer's characteristics and needs.

¹⁶ For a more extensive discussion, see RAND Institute for Civil Justice (2008).

¹⁷ *Fin. Planning Ass'n v SEC*, 375 U.S. App. D.C. 389, 2007.

¹⁸ Wrona (2012) argues that interpretations of the suitability standard and other FINRA rules impose detailed and rigorous standards of conduct on broker-dealers.

Broker-dealers have entered the advice business either by relying on the ‘solely incidental’ exemption of the Investment Advisers Act or by becoming dually registered as investment advisers to provide fee-based advisory services. Broker-dealers generally provide advice focused on specific products or transactions and generally do not monitor a client’s financial position on an ongoing basis. They are compensated on a commission basis.

Neither investment advisers nor broker-dealers are regulated under ERISA, which applies to employment-based plans such as 401(k)s and defined benefit plans.

Proposals for Change

Both the SEC and the DOL are considering changes that would affect broker-dealers.

SEC. The SEC proposal grows out of the 2010 Dodd-Frank “Wall Street Reform and Consumer Protection Act,” which required a study to assess the adequacy of the existing legal standards for personalized investment advice.¹⁹ The study recommended the adoption of a uniform fiduciary standard for investment advisers and broker-dealers that would require them, when dealing with retail customers, to act in the best interest of the customer without regard to their own financial interest.²⁰ The industry acknowledges that the cost of demonstrating that their recommendations are not merely suitable but in their client’s best interests would be modest. Even under current law, “brokers must know their clients and only recommend products that are suitable based on a particular client’s needs, time horizon, and risk profile” (Insured Retirement Institute 2011).²¹

DOL. The DOL proposed changes that would result in broker-dealers being classified as “fiduciaries” under the Internal Revenue Code when providing investment advice for IRAs. As such, they would be subject to IRS prohibited transaction rules (see Figure 4). Specifically,

¹⁹ An earlier draft of the Dodd-Frank legislation would have eliminated the broker-dealer exemption from the definition of investment adviser under the Investment Advisers Act of 1940.

²⁰ The extension of fiduciary status under ERISA and the Internal Revenue Code does not in itself result in fiduciary status under the securities law or coverage under the Investment Advisers Act of 1940.

²¹ Wrona (2012) argues that the requirements on broker-dealers are detailed and rigorous, under interpretations of the suitability standard and other FINRA rules, and in some respects may be more demanding than the requirements for investment advisers.

under this anti-self-dealing provision, the proposal would prevent broker-dealers from receiving third-party payments, such as 12b-1 fees. The prohibition on 12b-1 fees would apply only to the approximately 20 percent of total mutual fund assets held in IRAs.²² Several points are important here, because the industry reaction suggests considerable misunderstanding.

First, the DOL proposal does not change the general standard of conduct required of broker-dealers. They can continue to operate under a suitability standard rather than the “solely in the interest” standard required of ERISA fiduciaries. Thus, studies showing the additional hours required to satisfy the higher standard (for example, Oliver Wyman 2011) appear to be in error, since the standard of conduct would not change.

Second, substantial confusion appears around the prohibition of commission payments.²³ Here it is useful to distinguish between two types of commissions. The first is transactional commissions that broker-dealers receive for the purchase or sale of stocks, bonds or mutual funds. The second is ongoing payments from mutual funds. At this point, the DOL proposal would subject brokers who advise IRAs to a prohibition against only the second type of commission, which the agency fears could incent broker-dealers to sell clients excessively expensive products.

Third, the mechanism through which the DOL can make this change is the agency’s ability to define who is a fiduciary under ERISA and/or the tax code by reason of giving advice.²⁴ In 1975, the agency issued a five-part test that said that people are providing investment advice if they: 1) make recommendations on investing in, purchasing or selling securities or other property, or give advice as to their value; 2) on a regular basis; 3) pursuant to a mutual understanding that the advice; 4) will serve as a primary basis for investment decisions; and 5) will be individualized to the particular needs of the plan. Many broker-dealers avoid fiduciary

²² In 2011, total mutual fund assets amounted to \$11.6 trillion, and mutual funds in IRAs amounted to \$2.2 trillion (Investment Company Institute 2012).

²³ See, for example, American Benefits Council (2011); Insured Retirement Institute (2011); and Investment Company Institute (2011).

²⁴ ERISA gives DOL authority to define “fiduciary” for ERISA purposes. This authority was extended to IRAs under a subsequent reorganization plan that divided up responsibilities (Office of the President of the United States 1978).

status by including disclaimers in their written agreements with IRA holders, for example explaining that their advice will not constitute a “primary basis” for the IRA holder’s decisions or will not meet some other prong of the test.²⁵

The DOL proposal would sweep more broker-dealers into fiduciary status under the Internal Revenue Code by replacing the five-part test with a new test. Under the new definition, a person would be a fiduciary if he performs one of the following activities for a fee: 1) appraisals or fairness opinions concerning the value of securities or other property; 2) recommendations as to the advisability of investing in, purchasing, holding or selling securities or other property; or 3) recommendations as to the management of securities or other property, and if he meets one of four specified conditions: 1) represents to a plan, participant or beneficiary that he is acting as an ERISA fiduciary; 2) already functions as an ERISA fiduciary to the plan by virtue of having any control over the management or disposition of plan assets, or by having discretionary authority over the administration of the plan; 3) operates as an investment adviser under the Investment Advisers Act of 1940; or 4) provides the advice pursuant to an agreement or understanding that the advice may be considered in connection with investment or management decisions with respect to plan assets and will be individualized to the needs of the plan.

Assessing the Impact

This section makes an educated guess at the potential impact of the DOL proposal and then, as a consistency check, looks at the expectations for a similar, though more radical, approach in the United Kingdom.

A Likely Outcome for the United States

The most important aspect of the DOL proposal is that, with respect to IRA transactions, it would prohibit broker-dealers who give investment advice from receiving payments from mutual

²⁵ Their written material often includes phrases like “although we meet with you to discuss your needs, the advice is not individualized,” or “the assistance is educational in nature and should not be interpreted as recommendations; you should make your own decision,” etc.

funds. (It would also prohibit revenue sharing from other parties, such as advisers to mutual funds, and selling out of the broker-dealer's own inventory.) The goal is to eliminate incentive payments to broker-dealers for selling IRA customers high-cost funds. These payments are primarily 12b-1 fees, which amount to 25 basis points or less for no-load funds and are paid to the broker-dealer for as long as the customer holds the shares.²⁶ Most likely, mutual fund producers will continue to make the payments because IRAs are only a small portion (about 20 percent) of the mutual fund market, in which case the broker-dealer will be required to rebate the payment to the customer. If nothing else changes, customers will see their fees decline by the amount of the 12b-1 fees. 12b-1 fees were reported to amount to \$9.5 billion for all mutual funds in 2009, a year when mutual fund assets were roughly the same as 2011.²⁷ Assuming that the share of fees attributable to IRA customers is 20 percent, they should expect to receive rebates of about \$2 billion.

In addition to receiving rebates, IRA customers would gain from the DOL proposal to the extent that broker-dealers have less incentive to put them into high-fee funds. Since the high-fee funds tend to be actively managed (see Figure 5) and an extensive academic literature has documented that managed funds – at least in the equity arena – do not outperform the benchmarks, IRA holders would benefit by being directed to index funds, in which returns tend to be higher and fees significantly lower.²⁸

²⁶ 12b-1 fees are explicitly limited to 25 basis points or less for funds identifying themselves as “no-load” funds. This restriction does not apply to load funds, whose 12b-1 fees are capped at 100 basis points. For further details, see Investment Company Institute (2004).

²⁷ Schapiro (2010).

²⁸ A substantial literature analyzes whether mutual funds outperform benchmark indices before or after fees. A consistent finding is that, net of expenses, actively-managed funds underperform relevant indices. But the magnitude of the underperformance is sensitive to the period covered, the model specification, the choice of index, and the use of controls to correct for survivor bias. Malkiel (1995) concluded that, over the period 1982-1991, the average general equity mutual fund underperformed the S&P500 index by 183 basis points. Gruber (1996) estimated the average underperformance of actively-managed equity mutual funds at 194 basis points over the period 1985-1994, but this dropped to 65 basis points after controlling for fund characteristics. Malkiel (2005) estimated the underperformance of equity funds at 224 and 252 basis points over the 29 and 10 years ending December 31, 2003. In contrast, Wermers (2000) found that, over the period 1975-1994, mutual funds outperformed the CRSP value weighted index (a broad market index) by 130 basis points before expenses, but underperformed by 100 basis points after expenses. Using data from 1984 to 2006, Fama and French (2010) concluded that pre-expense returns are close to those of the market portfolio, and that, net of expenses, funds underperform the market by 81 to 100 basis points.

An open question is whether broker-dealers would take actions to offset the loss of the \$2 billion. The first step is to put the \$2 billion in perspective. According to the SEC, total broker-dealer revenues in 2010 amounted to \$247.8 billion (see Table 3).²⁹ Trading gains and profits from underwriting, which amounted to \$48.4 billion, could be subtracted to get a better measure of income from managing other people's money. This adjustment produces a figure of about \$200 billion. Thus, the loss of 12b-1 fees for mutual funds in IRAs (\$2 billion) would amount to about 1 percent of their total (non-trading/non-underwriting) annual revenue (\$200 billion). Broker-dealers could respond in a number of ways. First, they could simply accept lower profits. Second, they could make up the loss by increasing the price of transactional commissions, although such an increase might drive some customers to discount brokers. Third, they could raise their volume of transactional commissions by increased buying and selling of securities. Or fourth, they could institute platform charges; such charges are not prohibited because they are paid by the client and cover account administration.

Alternatively, broker-dealers could make up the \$2 billion by shifting to fee-based advisory accounts. Critics of the DOL proposal have argued that charging a percent of assets under management would result in small investors facing substantial additional costs, and in many households forgoing advice (Oliver Wyman 2011). The obvious question is whether broker-dealers are going to change what has been viewed as a successful business model in response to a 1-percent decline in revenues. It seems unlikely. But if broker-dealers did change their business model, the outcome could be positive or negative.

In most markets, transparency of charges is thought to benefit consumers because it enables them to base their consumption decisions on a proper understanding of the cost of the product. The concern of the industry must be that households may not be willing to pay for financial advice. This problem is not unique to financial services; in a market economy, sellers are expected to make the business case for purchase of their product or service.

A second concern is that, if broker-dealers were to adopt a fee-based approach, charges may

²⁹ U.S. Securities and Exchange Commission (2012).

increase. According to industry advocates, typical asset management fees exceed typical commissions (Oliver Wyman 2011).³⁰ It is not clear, however, whether this alleged differential reflects differences in costs of provision, scope of services, or in willingness to pay. If the latter explanation is more correct, then a simple comparison of the two models will overstate the additional cost. Alternatively, asset-based pricing may be more expensive if it is associated with a higher quality of care. Although the proposed DOL rule does not impose a “sole interest” fiduciary standard, broker-dealers who switched from transaction to asset-based charging would likely be subject to the Investment Advisers Act of 1940, which does require a full fiduciary standard. Even so, advisers already gather much of the information required to make “sole interest” recommendations, so the shift might result in only modest additional costs (Insured Retirement Institute 2011).

Thus, the best prediction of the impact of the DOL proposal is that broker-dealers will not take any offsetting actions and fees will decline. The minimum decline would be the rebate of the 12b-1 charges, about 4 basis points on total IRA assets.³¹ Any additional reduction in fees depends on the extent to which broker-dealers steer customers away from actively-managed mutual funds (with expense ratios of 93 basis points for equity funds and 66 basis points for bond funds, including 12b-1 fees) and towards index funds (14 basis points). If one third of mutual fund assets (including both equity and bond funds) were shifted to low-fee index funds, total IRA fees could fall by another 7 basis points.³² Additionally, if one believes the estimates

³⁰ The Wyman report omitted indirect fees like 12-b1. In response to a DOL request, they later provided this piece, which narrowed the gap between asset-based and commission (plus indirect) fee-based costs. Still, it is unclear whether full service brokerage is less expensive than asset-based advisory account charges, since Wyman included both full service and discount brokerage without any breakdown. Moreover, Wyman assumes that if brokers were to charge directly, they would charge prices as high as investment advisers even though they would be providing more limited services.

³¹ Total IRA assets amount to \$5,126 billion, so IRA 12b-1 fees of \$2 billion amount to about 0.04 percent of assets, or 4 basis points.

³² The 7-basis point estimate is derived as follows. The first step is to estimate the percentage of mutual funds levying 12b-1 fees. If these fees average 25 basis points, and 45 percent of IRA assets is invested in mutual funds (as reported in Investment Company Institute 2012), an estimate of \$2 billion of IRA 12b-1 fees is consistent with one-third of IRA mutual funds levying such fees. Using data from the Investment Company Institute (2012), we then assume that fees are reduced on one third of the 45 percent of IRA assets invested in mutual funds. We further assume that 64 percent of such assets are invested in stock funds, and the remaining 36 percent in bond funds, and that fees on stock funds decline from 93 to 14 basis points, and fees on bond funds from 66 to 13 basis points. The final step is to subtract the assumed saving in 12b-1 fees.

that actively-managed equity funds underperform index funds by as much as 224 basis points as reported in Malkiel (2005), then a shift of one third of IRA mutual fund assets to index funds could produce another 13 basis points, but that should be considered a maximum.³³ A reduction in fees between 4 basis points and 24 (4+7+13) basis points would save the consumer between \$2 billion and \$12 billion. This modest gain for consumers is unlikely to cause a meaningful disruption in the provision of advice.

Speculation from the United Kingdom

In order to assess whether such an outcome is a reasonable prediction for the U.S. DOL proposal, it is useful to look at the U.K.'s Retail Distribution Review (RDR) program, which was launched in 2006 in response to mis-selling scandals.³⁴ The program, which comes into full effect at the end of 2012, has three objectives: 1) to create a transparent system for charging for financial advice; 2) to ensure that consumers understand whether the advice they are receiving is independent or restricted; and 3) to improve professional standards and knowledge.³⁵ To achieve these objectives, the program: 1) prohibits financial advisers from receiving all but transaction commissions; 2) requires advisers to disclose whether they are independent, providing unbiased and unrestricted advice; and 3) requires advisers to subscribe to a code of ethics, hold appropriate qualifications, and undertake continuing professional development.³⁶

The RDR program is considerably more extensive than the DOL proposal for a number of reasons. First, it affects both retirement and non-retirement accounts. As noted earlier, in the

³³ The 13 basis points equals additional underperformance of 224 minus 93 basis points, multiplied by the 45 percent of IRA assets invested in mutual funds, the 64 percent of IRA mutual funds invested in stocks, and the one third that is assumed to switch from underperforming actively-managed funds to index funds. The calculations further assume that bond mutual funds do not underperform relevant indices, after accounting for fees.

³⁴ Between 1988 and 1994, more than three million people were incorrectly advised to take out personal pension plans when they would have been better off in their employer scheme. This situation was followed by a scandal in which as many as five million households were sold interest-only mortgages, with repayment to be effected through an endowment policy, being incorrectly informed that their policy was guaranteed to pay off their mortgage. Between 2000 and 2002, about 25,000 individuals lost money in a split capital investment trust scandal, losing as much as £600 million. Finally, as many as 12 million individuals may have been mis-sold payment protection insurance over the past 10 years, leading banks to reserve £7.5 billion.

³⁵ Financial Services Authority (2011).

³⁶ Complementing the above reforms is an attempt to increase household "financial capacity." The Money Advice Service, established in 2010 as the Consumer Financial Education Body, is responsible for enhancing the public's understanding and knowledge of financial matters and their ability to manage their financial affairs.

U.S., IRAs hold only 19 percent of mutual funds. Second, total expense ratios on mutual funds are larger in the United Kingdom than in the United States.³⁷ Third, starting after 12/31/12, the program *requires* that customers are charged explicitly for advice, whereas the DOL proposal has no such requirement. Fourth, the United Kingdom requires a comprehensive and fair analysis of the relevant market, a higher standard than the suitability standard currently required for broker-dealers in the United States.³⁸

The question is how the RDR is expected to affect fees.³⁹ According to experts at the U.K.'s Financial Services Authority (FSA), prior to the RDR, annual management charges on actively-managed unit trusts have averaged around 150 basis points. Post-RDR, the FSA estimates that product manufacturers initially will charge a price of about 75 basis points. An increasing proportion of investors use a "platform" to invest in a portfolio of mutual funds, which will charge 25 basis points. At these cost levels, the fee will decline by 50 basis points, before accounting for the cost of advice, which will be charged separately.

The projected post-RDR fees are still quite high. Surveys suggest that the cost of index funds can be as low as 10 basis points, and those of globally-diversified, actively-managed portfolios can be as low as 51 basis points.⁴⁰ Platforms are available for close to zero out-of-pocket cost.⁴¹ If advisers are required to undertake a "comprehensive and fair analysis of the relevant market," it is hard to see how product and platform costs of 75 and 25 basis points, respectively, can

³⁷ In the United Kingdom, mutual funds are referred to as "unit trusts." Khorana, Servaes, and Tufano (2009) report total sales charges on U.K. and U.S. equity funds at 248 and 153 basis points respectively, and total sales charges on bond funds at 173 and 105 basis points, respectively. Total sales charges include sales loads, where appropriate.

³⁸ Intermediaries are also permitted to provide "restricted advice," recommending products supplied by a single financial institution. An intermediary providing "restricted advice" must disclose the fact. He will not carry out a "comprehensive and fair analysis of the relevant market," but must ensure that any recommendations are "suitable" (see Financial Services Authority 2012).

³⁹ The focus of the analysis is the "fee wedge" for investment products. Data are not yet available on post-RDR insurance product prices, and many products have been redesigned in ways that render price comparisons difficult.

⁴⁰ For example, the Vanguard FTSE 100 ETF has a total expense ratio of 10 basis points. The Association of Investment Companies publishes data on the ongoing charges of closed-end investment funds. Ongoing charges for global growth and income funds average 73 basis points, but charges can be as low as 51 basis points.

⁴¹ Platforms currently receive rebates from unit trusts, which some pass on to their clients. These rebates will be prohibited by the RDR. Platforms are currently revising their pricing to reflect this prohibition. Offerings include on-line dealing charges as low as £9.95, with either a zero or nominal annual charge. Except for the smallest accounts, these charges would be far less than 25 basis points, consistent with our very low estimates of the costs of maintaining platforms.

survive. But financial advisers may continue to recommend high-cost, actively-managed unit trusts that they know and understand. Indeed, a recent financial report notes that publicly-traded unit trust companies have reported that they do not expect the RDR to adversely affect their profitability.⁴²

The other issue is the impact of the RDR on the cost and availability of advice. The RDR requires households to make an explicit payment for advice. The cost of that advice depends on hours required and hourly rates. The Association of British Insurers (2010) estimated the average time for the full advice process at 7.5 hours. Based on an hourly labor cost of £200 (Ernst and Young 2011), this estimate equates to a per-client cost of £1,500.⁴³ It is as yet unclear whether clients will be willing to pay that level of charges.⁴⁴ Some indication can be obtained from surveys of advisers' career plans. Atkin et al. (2011) show that just 8 percent of retail investment advisers reported planning to cease retail advice after 2012.⁴⁵ Those planning to quit included a disproportionate number of older advisers. An earlier study (Deloitte LLP 2009) found that only 21 percent of advisers thought they would be unable to move to adviser-based charging, with an additional 33 percent being unsure.⁴⁶ The findings suggest that the majority of U.K. advisers think that they will be able to sell their services. Thus, at least at this point, the U.K. reform is not anticipated to disrupt the availability of advice.

In short, reform proposals both here and in the United Kingdom involve a trade-off between the gain from lower fees and any adverse effect from the loss of advice. An overall assessment requires some measure of what consumers might lose if they did not have advice and how that loss compares to the effect of lower fees. The optimization model, described below, attempts to provide such magnitudes.

⁴² St James's Place Wealth Management (2012).

⁴³ Financial advisers spend considerable amounts of time developing sales leads. Post-RDR, instead of selling a financial product, advisers will be selling a financial plan that may or may not include the use of financial products. But they will still be selling, and must still devote time to prospecting. The 70-percent utilization rate assumed by Ernst and Young (2011) may be too high, resulting in an understatement of costs.

⁴⁴ Charges can be paid out of invested assets, making them less visible than they would be if households were required to write a check.

⁴⁵ Oxera (2009).

⁴⁶ Deloitte LLP (2009).

The Optimization Model

Any U.S. reform will likely reduce fees to some extent, thereby improving net returns for investors. Bolder reforms will likely yield larger improvements. With higher net returns, households can save less to achieve any given level of retirement income and will be able to enjoy higher consumption both before and after retirement. The purported risk is that reform may cause broker-dealers to limit the availability of financial advice, resulting in households choosing inappropriate asset allocations, failing to diversify their portfolios, and failing to save enough for retirement.

The optimization model is first used to calculate risk-adjusted measures of the impact of excessive fees and the risk of making financial mistakes on lifetime consumption for investors holding *all* of their financial assets in IRAs. The impact on investors with only part of their financial assets in IRAs will be proportionately less. Although reforms are unlikely to significantly reduce the availability of financial advice and thus impose offsetting costs, the above calculations are then used to investigate the terms of a potential trade-off between lower fees and poorer financial decision-making. It shows that although it is possible to construct scenarios in which costs exceed benefits, those scenarios require quite extreme assumptions about the percent of households forgoing advice and the types of mistakes made by those that forgo advice.

The exercise starts by calculating optimal behavior for a household facing mortality risk and labor market uncertainty.⁴⁷ Data limitations and the limits of computational feasibility require simplifying assumptions. The model is described in more detail in the appendix. The goal of the household is to maximize expected discounted lifetime utility. In each period, the household chooses how much to save and how to allocate its portfolio between stocks and a risk-free

⁴⁷ The logarithmic mean and standard deviation of real stock returns are assumed to be 6.5 percent and 20 percent, respectively. Bonds are assumed to yield a 3-percent, risk-free real return. In reality, the real yield on long-dated bonds fluctuates considerably, while there is also considerable evidence that stock returns exhibit mean reversion. Campbell and Viceira (2002) show that long-term bonds are the true risk-free asset for long-term investors because they provide a guaranteed income. If stock returns exhibit mean reversion, then the optimal equity allocation should reflect whether expected returns are high or low relative to the long-run average. But incorporating the above features would greatly complicate our model, without yielding any important additional insights.

bond.⁴⁸ The household chooses the market portfolio that offers an optimal trade-off between risk and return. The household participates in Social Security. It invests in stock and bond index mutual funds with expense ratios of 14 and 13 basis points, respectively.⁴⁹ The model assumes a coefficient of risk-aversion of five.

Benefits of Lower Fees

The exercise first considers the potential benefits from reducing fees, measured as the percentage of salary that a household would pay to avoid a high-fee investment. That is, it calculates the percentage of salary that would leave the household indifferent between the optimal portfolio and a higher-fee portfolio.⁵⁰ It assumes that households change neither their portfolio allocation nor their saving rate in response to the decline in net returns, so the calculations represent an upper-bound estimate of required compensation.

The model considers three scenarios: 1) 12b-1 fees are eliminated, resulting in an increase in net-of-fee returns of 4 basis points; 2) the returns on actively-managed stock and bond mutual funds are reduced by fees in excess of those for low-cost index funds; and 3) the returns of actively-managed *equity* funds are reduced by the excess fees *and* by underperformance relative to the relevant indices as reported in the literature. The model assumes that actively-managed stock and bond funds have expense ratios of 93 and 66 basis points, respectively.⁵¹ Estimates of the extent to which the average actively-managed stock mutual fund underperforms relevant market indices vary considerably. The calculations assume, based on Malkiel (2005), an upper bound estimate of 224 basis points (including the impact of fees).

⁴⁸ The model does not distinguish between 401(k), IRA, and non-retirement financial assets. Most households hold little financial wealth outside of their 401(k) and IRA.

⁴⁹ Investment Company Institute (2012).

⁵⁰ This approach is analogous to the annuity equivalent wealth calculation made in the annuitization literature (Brown and Poterba 2000), the percentage increase in age 65 wealth that would leave the household indifferent between an actuarially fair annuity and an optimal decumulation of unannuitized wealth. One cannot simply compare age 65 wealth under the baseline and alternative scenarios because households will optimally choose to enjoy part of the benefit of higher net of expense returns in the form of higher pre-retirement consumption.

⁵¹ Investment Company Institute (2012). Although a substantial literature has investigated the trading costs and underperformance of stock funds, there has been no comparable research into bond funds. We therefore conservatively assume zero trading costs or underperformance for such funds.

The results are reported in Table 4. A household investing in actively-managed funds that did not levy 12b-1 fees would pay 0.3 percent of salary to avoid such fees. A household investing in low-cost index stock and bond funds would pay 6.6 percent to avoid investing in actively-managed funds and 10.0 percent when excess fees and underperformance are taken into account. These amounts closely approximate the percentage impact of higher fees and underperformance on the expected present value of lifetime consumption.

Costs of Investor Mistakes

The model then considers three types of mistakes that investors might make: 1) choosing an inappropriate asset allocation; 2) failing to diversify the stock component of their portfolios⁵² so that, while they enjoy the same expected return, they take on uncompensated risk; and 3) saving too little. For (1), it assumes that the household invests too much or too little of its portfolio in stocks at all ages. For (2), it assumes that instead of holding a diversified stock portfolio, the household holds a portfolio of ten, four, or two stocks, with corresponding increases in the standard deviation of the return on its stock portfolio (Statman 1987). For (3), it assumes that a household saves nothing at all. It is assumed that two thirds of IRA investors are in either an advisory relationship or hold a discount brokerage account. It is only the remaining one third who are at risk of forgoing advice. Two scenarios are considered. The first, and most likely, is that the reform has no effect on the availability of advice. The second assumes that 50 percent of the households “at risk” forgo advice and then make one of the above investment mistakes.

Table 5 reports the percentages of salary that households would be willing to pay to avoid making the above mistakes, averaged over both those who forgo advice and those whose use of advice is unaffected by the reform. If zero percent forgo advice, the cost of the reform is zero. But even if an implausible 50 percent of the one third of IRA holders assumed to be “at risk” forgo advice, the impact of substantial asset allocation mistakes is relatively modest. The costs of a 100 percent and a zero percent asset allocation to stocks at all ages amounts to at most 0.42 and 0.63 percent of salary, respectively, and the costs of failing to diversify one’s portfolio

⁵² Estimates of the standard deviations of undiversified portfolios are taken from Statman (1987).

amount to at most 0.65 to 0.98 percent of salary, and in each case could be as low as zero. These at most moderate effects reflect the fact that only a minority of households is even at risk of forgoing advice, and the insurance provided by labor market earnings, Social Security and, for undiversified stock portfolios, the household's bond holdings.

The calculations in Tables 4 and 5 can then be used to evaluate the net effects of reducing fees. The benefit of eliminating 12b-1 fees is equivalent to 0.3 percent of salary. The cost depends on the percent of households that forgo advice and the severity of the resulting investment mistakes they make. Our best estimate is that a 4-basis-point reduction in fees will have no discernible effect on the supply of financial advice.

Larger reductions in fees would, of course, bring larger benefits but might also result in some reductions in the supply of financial advice. Nevertheless, in such a scenario, the benefits would still dwarf the costs. For example, as noted in Table 4, shifting households to index funds would substantially increase the benefits from lower fees – to 6 percent or more of salary. On the cost side, even under the extreme assumption⁵³ that these lower fees result in half of households forgoing advice, the estimates of the individual investment mistakes are all 1 percent of salary or lower.

Bolder Proposals

Although the optimization model shows that the benefits of a DOL-type proposal resulting from lower fees exceed any likely costs resulting from financial mismanagement due to the lack of advice, the impact of the proposal is modest. Basically, the gain to the consumer is a guaranteed 4 basis points due to the rebate of 12b-1 commissions. Consumers could see a potential additional gain of 7 basis points *if* broker-dealers redirect one third of mutual fund holdings from high-fee, actively-managed mutual funds to low-fee index funds. And *if* one believes the estimates that actively-managed funds underperform index funds, even before fees, consumers

⁵³ This assumption is extreme because it requires implausible assumptions about the price elasticity of the supply of financial advice.

could see a further gain of perhaps 13 basis points. But the 4 basis points is all that policymakers can really count on.

Given the potentially modest impact of the DOL proposal, more ambitious reforms to reduce investment fees merit consideration. The policy justification for such reforms is that, given the tax advantages provided to 401(k)s and IRAs, the government needs to ensure that the accounts are managed solely in the interests of participants. High fees frustrate this policy objective. The proposed options fall into four categories: 1) making it easier to retain accumulations within the 401(k) system; 2) making the rollover from a 401(k) to an IRA an ERISA-covered event; 3) extending ERISA to all rollover IRAs; and 4) instituting changes to further control fees in both 401(k)s and IRAs. Some of these changes could be accomplished through rule-making, while others would require legislation.

Making It Easier to Keep Money in 401(k)s

At a minimum, participants should be encouraged to keep their money in the 401(k) system when switching jobs, rather than rolling balances over into IRAs. Keeping money in 401(k)s has three advantages: 1) 401(k)s are covered by ERISA fiduciary standards, which require financial advisers to act solely in the participant's interests; 2) recent DOL disclosure requirements have helped shine a spotlight on fees; and 3) 401(k)s operate in a wholesale environment, lending them potential pricing advantages in dealing with investment managers.

Two straightforward changes could help here. First, workers switching jobs should always be allowed to keep their 401(k) assets with their previous employer. This proposal would require a change in the provision that allows employers to cash out account balances of less than \$5,000.⁵⁴ Second, workers should always be allowed to move their 401(k) assets from a previous employer to a new employer. This proposal would require a change in the provision that gives employers the option to deny a rollover from a previous plan.⁵⁵ Conversations with experts suggest that employers would not be opposed to retaining accounts or accepting accounts from former

⁵⁴ IRS Code 411(a)11 and ERISA 203(e).

⁵⁵ IRS Regulation 1.401(a)(31)-1, A-13.

employers, because higher balances give them more leverage when negotiating fees. Both changes would increase the likelihood that participants stay in the relatively protected 401(k) environment.

Regulating Rollover Transactions

Even with changes to keep money in the 401(k) system, some will want to roll over their balances to IRAs. In this case, it is important that they think carefully before moving their money. One option is to make any rollover transaction subject to ERISA, given that the assets in 401(k)s come from the employer plan arena.⁵⁶ Such a change would mean that an adviser could recommend a rollover only when it was solely in the client's interests, as the adviser would be subject to the higher standard required of 401(k) fiduciaries. Participants considering a rollover could also be presented with disclosure forms comparing fees in their 401(k) plan with those in their proposed IRA and showing the respective impacts on projected wealth at retirement. Finally, if a 401(k) participant does decide to go ahead with an IRA rollover, policymakers could set a default investment vehicle of a life-cycle index fund.

Extending ERISA to All Rollover IRAs

The most sweeping reform option would be to extend ERISA protections to all rollover IRAs. The rationale is that rollover money has been accumulated in the employer plan arena, which is protected by ERISA's fiduciary standards and fee disclosure, and that the concern for protecting these funds is not lessened by their movement into another form of account. Most likely, if the enactors of ERISA had envisioned that most defined contribution money would end up in IRAs, they would have ensured ERISA-type protections for these accounts. The change might create its own complications, requiring other modifications to ERISA and creating new overlaps in agency jurisdictions.

⁵⁶ The DOL might be able to accomplish this change by regulation. If legislation is required, it would involve amending ERISA Title 1, Section 4(a) to include coverage of rollover IRAs.

Controlling Fees

The DOL has undertaken a major effort to ensure that employees have access to low-cost funds, but four additional options would greatly improve the fee situation. These options include: establishing benchmarks for 401(k) fees; requiring reporting and benchmarks for IRA fees; requiring 401(k) plans to offer index funds; and eliminating high-cost, actively-managed funds.

First, existing 401(k) fee disclosures could be enhanced.⁵⁷ Recipients of the current fee disclosures receive information on the fees they are paying, but have no clear benchmark against which to determine whether these fees are excessive. An option here would be to require the disclosure form to compare the costs of the individuals' current investments with those of the typical stock or bond index fund, along with an estimate of the percentage increase in wealth at age 65 from switching to the index fund.⁵⁸

Second, providers of IRAs could be required to report on the asset holdings and fees charged in these accounts. This information would make it possible for people considering rolling over their balances to compare the 401(k) fees with those in IRAs.

Third, all 401(k) plans could be required to offer low-cost index funds, including an equity fund, a bond fund, and a life-cycle fund. As part of this proposal, the government could give a "seal of approval" to low-cost funds that meet certain criteria.

Finally, a more ambitious reform is to limit investment options to low-cost index funds for 401(k)s and, if ERISA were extended, for rollover IRAs as well. As discussed earlier, virtually all researchers agree that most actively-managed equity funds can be expected to underperform index funds once fees are considered. It makes no sense to expose the average participant to

⁵⁷ The current fee disclosure requirements have only been in place a short time, so it is not possible to fully gauge their effects. One recent survey (Plan Sponsor Council of America 2012) reported that the requirements have so far had little direct effect. However, anecdotal evidence suggests that some fund providers have already begun to introduce lower fee versions of some of their funds. And the impact of the requirements could increase over time as participants and plan sponsors have more exposure to the new disclosure data.

⁵⁸ Small 401(k) plans are more expensive to administer, a concern that could be addressed through an appropriate disclosure on the fee statement.

these options. If people want to buy actively-managed funds with their non-tax-advantaged saving, that is fine. But in plans that cost the taxpayer money, investing should be cost effective. A variant of such a proposal would leave some room for actively-managed funds with low fees.

In short, a number of options are available for controlling fees beyond those already implemented by the DOL.

Conclusion

Additional protections are required in the IRA market. As long as accumulations are held in 401(k) plans, participants remain in a world in which sponsors must operate as fiduciaries and fees are under a spotlight. Once they roll over their accounts into IRAs, they enter a world where suitability becomes the standard of care and broker-dealers are paid commissions that encourage the sale of high-priced mutual funds. If a fiduciary standard and attention to fees are appropriate for retirement assets when they are in a 401(k) plan, then such safeguards are clearly still appropriate when they are rolled over. The SEC and the DOL proposals reflect this logic.

Although the DOL proposal has met with a storm of controversy, it involves a modest change in the form of eliminating 12b-1 and other fees that might incent broker-dealers to misdirect their clients' investments to high-fee products. In the short run, the direct impact of such a change would be rebates to IRA investors of about 4 basis points. Gains could be greater if broker-dealers responded by shifting investments to low-cost index funds. The purported industry concern is that, under the DOL proposal, low- and middle- income households would lose their access to financial advice and make costly mistakes that would reduce their holdings at retirement. Such an outcome seems unlikely for two reasons. First, broker-dealers are unlikely to change their business model in response to a 1-percent reduction in non-trading revenues. Second, even if, in the long run, some IRA holders lost advice as a result of a move to lower-fee funds, such mistakes would have to be both widespread and egregious to offset the gain from lower fees.

Given that the DOL proposal is likely to have only a small impact, it is worth considering bolder and more direct approaches to controlling fees. These steps include: establishing benchmarks for

401(k) fees; requiring reporting and benchmarks for IRA fees; requiring 401(k) plans to offer index funds; and eliminating high-cost actively-managed funds. In addition, broader changes, such as making it easier to retain accumulations within the 401(k) system; making the rollover from a 401(k) to an IRA an ERISA-covered event; and extending ERISA to all rollover IRAs, would greatly enhance protections.

In short, the DOL proposal has highlighted an important issue – namely, the enormous growth in rollover IRAs – but it should be viewed as only a modest first step.

References

- American Benefits Council. 2011. "Comments on RIN 1210-AB32, Definition of Fiduciary Proposed Rule." (February 3). Washington, DC.
- Ameriks, John and Stephen P. Zeldes. 2001. "How Do Household Portfolio Shares Vary with Age?" Working Paper 6-120101. New York, NY: TIAA-CREF Research Institute.
- Association of British Insurers. 2010. "Increasing Consumer Access to Advice." London, UK.
- Atkin, Bryan, Naomi Crowther, Dominika Wintersgill, and Andrew Wood. 2011. "Progress Towards the Professionalism Requirements of the Retail Distribution Review." RS Consulting Report.
- Benitez-Silva, Dwyer, D.S., Heiland, F., and Sanderson, W.C. 2007. "Retirement and Social Security Reform Expectations: A Solution to the New Early Retirement Puzzle." Unpublished Working Paper.
- Bergstresser, Daniel, John Chalmers, and Peter Tufano. 2009. "Assessing the Costs and Benefits of Brokers in the Mutual Fund Industry." *The Review of Financial Studies*.
- Bromberg, Jason and Alicia P. Cackley. 2012. "Regulating Financial Planners: Assessing the Current System and Some Alternatives." Working Paper 2012-18. Philadelphia, PA: Pension Research Council.
- Brown, Jeffrey R. and James M. Poterba. 2000. Joint Life Annuities and the Demand for Annuities by Married Couples. *The Journal of Risk and Insurance* 67(4): 527-553.
- Campbell, John Y. and Luis M. Viceira. 2002. *Strategic Asset Allocation: Portfolio Choice for Long-Term Investors*. Oxford University Press, Oxford, UK.
- Chalmers, John and Jonathan Reuter. 2012. "What is the Impact of Financial Advisors on Retirement Portfolio Choices and Outcomes?" Working Paper No. 18158. Cambridge, MA: National Bureau of Economic Research.
- Chetty, Raj. 2003. "A New Method of Estimating Risk Aversion." Working Paper 9988. Cambridge, MA: National Bureau of Economic Research.
- Chiappori, Pierre-Andre and Monica Paiella. 2011., Relative Risk Aversion Is Constant: Evidence from Panel Data. *Journal of the European Economic Association* 9(6): 1021-1052.
- Degeorge, Francois, Dirk Jenter, Alberto Moel, and Peter Tufano. 2004. "Selling Company Shares to Reluctant Employees: France Telecom's Experience. " *Journal of Financial Economics* 71(1): 169-202.

- Deloitte and Touche LLP. 2008. "Costing Intermediary Services: Financial Assessment of Investment Intermediaries." Report for the Financial Services Authority.
- Deloitte LLP. 2009. "Firm Behavior and Incremental Compliance Costs." Research for the Financial Services Authority.
- Ernst and Young. 2011. "RaDaR: Life and Pensions Outlook for 2011." London, UK.
- Fama, Eugene F., and Kenneth R. French. 2010. "Luck versus Skill in the Cross-Section of Mutual Fund Returns." *Journal of Finance* 55(5): 1915-1947.
- Financial Services Authority. 2011. "Distribution of Retail Investments: Delivering the RDR – Professionalism." London, UK. Available at: http://www.fsa.gov.uk/pubs/policy/ps11_01.pdf.
- Financial Services Authority. 2012. "Retail Distribution Review: Independent and Restricted Advice." Available at: <http://www.fsa.gov.uk/static/pubs/guidance/fg12-15.pdf>
- Fischel, Daniel R. and Todd D. Kendall. 2011. "Comment to the Department of Labor on a Proposed Rule Regarding Fiduciary Status under ERISA." Chicago, IL: Compass Lexecon. Available at: <http://www.dol.gov/ebsa/pdf/1210-AB32-PH056.pdf>
- Forbes. 2011. "12b-1 Fees." (July 12). New York, NY. Available at: <http://www.forbes.com/sites/forbesfinancialglossary/2011/07/12/12b-1-fees>.
- Frankel, Tamar. 2011. "The Regulation of Brokers, Dealers, Advisors and Financial Planners." *Review of Banking and Financial Law* (Boston University School of Law) 30 (2010-11): 123-139.
- Gil-Bazo, J. and Ruiz-Verdu, P. 2009. "The Relation between Price and Performance in the Mutual Fund Industry." *Journal of Finance*. 64(5): 2,153-2,183.
- Gil-Bazo, J. and Ruiz-Verdu, P. 2008. "When Cheaper is Better: Fee Determination in the Market for Equity Mutual Funds." *Journal of Economic Behavior and Organization*. 67(3-4): 871-885.
- Gruber, Martin J., 1996. "Another Puzzle: The Growth in Actively Managed Mutual Funds." *Journal of Finance* 51: 783–810.
- HM Treasury. 1998. "CAT Standards for Individual Savings Accounts." (December). London, UK. Available at: www.fsa.gov.uk/pubs/additional/pia_cat.pdf.
- Insured Retirement Institute. 2011. Testimony of Theresa M. Atanasio to the Employee Benefits Security Administration. (March 1). Washington, DC.

- Internal Revenue Service. 2012. "Social Security and Railroad Equivalent Retirement Benefits." Publication 915. Washington DC: U.S. Department of the Treasury.
- Investment Company Institute. 2012. "2012 Investment Company Fact Book." Washington, DC.
- Investment Company Institute. 2011. "Comments on the Department of Labor Public Hearing on Definition of Fiduciary." (April 12). Washington, DC.
- Investment Company Institute. 2004. "The Cost of Buying and Owning Mutual Funds." Research in Brief, Volume 13, No. 1. Washington, DC.
- Khorana, Ajay, Henri Servaes, and Peter Tufano. 2009. "Mutual Fund Fees Around the World." *Review of Financial Studies* 22(3): 1,279-1,310.
- Kopcke, Richard W., Francis M. Vitagliano, and Zhenya S. Karamcheva. 2009. "Fees and Trading Costs of Equity Mutual Funds in 401(k) Plans and Potential Savings from ETFs and Commingled Trusts." Working Paper 2009-27. Chestnut Hill, MA: Center for Retirement Research at Boston College.
- Madrian, Brigitte C. and Dennis F. Shea. 2001. "The Power of Suggestion: Inertia in 401(k) Participation and Savings Behavior." *Quarterly Journal of Economics* 116(4): 1149-87.
- Malkiel, Burton G. 1995. "Returns from Investing in Equity Mutual Funds 1971 to 1991" *Journal of Finance* 1(2): 549-572.
- Malkiel, Burton G. 2005. "Reflections on the Efficient Market Hypothesis: 30 Years Later." *The Financial Review* 40: 1-9.
- Maurer, Raimond, Olivia S. Mitchell, and Ralph Rogalla. 2010. "The Effect of Uncertain Labor Income and Social Security on Lifecycle Portfolios." In *Reorienting Retirement Risk Management*. Oxford, United Kingdom: Oxford University Press.
- Mullainathan, Sendhil, Markus Noeth, and Antoinette Schoar. 2012. "The Market for Financial Advice: An Audit Study" Working Paper No. 17929. Cambridge, MA: National Bureau of Economic Research.
- Munnell, Alicia H. and Annika Sundén. 2004. *Coming Up Short: The Challenge of 401(k) Plans*. Washington, DC: Brookings Institution Press.
- Office of the President of the United States. 1978. "Executive Order: Reorganization Plan No. 4 of 1978." Washington, DC.
- Oliver Wyman. 2011. *Oliver Wyman Report: Assessment of the Impact of the Department of Labor's Proposed "Fiduciary" Definition Rule on IRA Consumers*. New York, NY: Oliver Wyman.

- Oxera. 2009. *Retail Distribution Review Proposals: Impact on Market Structure and Competition*. Oxford, UK.
- Paiella, Monica. 2001. "Limited Financial Market Participation: A Transaction Cost-Based Explanation." Working Paper W01/06. London, United Kingdom: Institute for Fiscal Studies.
- Plan Sponsor Council of America. 2012. "Snapshot Survey on Fee Disclosure." Chicago, IL.
- RAND Institute for Civil Justice. 2008. *Investor and Industry Perspectives on Investment Advisers and Broker-Dealers*. Technical Report.
- Schapiro, Mary L. 2010. "Opening Statement at the SEC Open Meeting – 12b-1." (July 21). Washington, DC: U.S. Securities and Exchange Commission."
- Scholz, John Karl, Ananth Seshadri, and Surachai Khitatrakun. 2006. "Are Americans Saving "Optimally" for Retirement?" *Journal of Political Economy* 114(4): 607-643.
- St. James's Place Wealth Management. 2012. "Investor Overview." Cirencester, UK. Available at: <http://www1.sjp.co.uk/investor-relations/investment-case/investor-overview.aspx>.
- Statman, Meir. 1987. "How Many Stocks Make a Diversified Portfolio?" *The Journal of Financial and Quantitative Analysis* 22(3): 353-363.
- Thaler, Richard H. and Shlomo Benartzi. 2004. "Save More Tomorrow Using Behavioral Economics to Increase Employee Saving." *Journal of Political Economy* 112 (1, pt. 2): S164-S187.
- U.S. Board of Governors of the Federal Reserve System. *Flow of Funds Accounts of the United States*, 2012 (December 6). Washington, DC.
- U.S. Board of Governors of the Federal Reserve System. *Survey of Consumer Finances*, 1983, 1995, and 2010. Washington, DC.
- U.S. Securities and Exchange Commission. 2011. *Study on Investment Advisers and Broker-Dealers*. Washington, DC. Available at: <http://www.sec.gov/news/studies/2011/913studyfinal.pdf>
- U.S. Securities and Exchange Commission. 2011. "Select SEC and Market Data Fiscal 2011." Washington, DC. Available at: <http://www.sec.gov/about/secstats2011.pdf>
- Vissing-Jorgensen, Annette. 2002. "Towards an Explanation of Household Portfolio Choice Heterogeneity: Nonfinancial Income and Participation Cost Structures." Working Paper 8884. Cambridge, MA: National Bureau of Economic Research.

Wermers, Russ. 2000. "Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transactions Costs, and Expenses." *Journal of Finance* 55(4): 1655-1695.

Wrona, James S. 2012. "The Best of Both Worlds: A Fact-Based Analysis of the Legal Obligations of Investment Advisers and Broker-Dealers and a Framework for Enhanced Investor Protection." *The Business Lawyer* 68(1): 1-56.

Young, Jean A. 2012. "Target-Date Fund Adoption in 2011." Research Note. Valley Forge, PA: Vanguard.

Table 1. *Number of Individuals with 401(k)s and IRAs, By Age Group, Millions*

Age	401(k) only	IRA only	401(k) & IRA	Total
20-29	4.1	1.2	0.4	5.7
30-39	8.3	3.3	2.3	14.0
40-49	9.8	4.8	3.4	18.0
50-59	7.8	7.7	4.7	20.3
60-69	2.3	7.1	2.1	11.6
70+	0.1	7.1	0.3	7.5
Total	32.5	31.3	13.3	77.1

Note: Total includes 20,000 individuals under age 20 that invest in a 401(k).

Source: U.S. Board of Governors of the Federal Reserve System, *Survey of Consumer Finances* (2010).

Table 2. *Number of Households Investing in 401(k)s and IRAs, by Income Quintile, Millions*

Quintile	401(k) only	IRA only	401(k) & IRA	Total
1	1.0	2.0	0.1	3.1
2	4.3	3.3	0.7	8.4
3	5.4	3.8	1.6	10.9
4	8.1	5.1	3.2	16.4
5	6.0	4.9	8.2	19.1
Total	24.9	19.2	13.8	57.9

Source: U.S. Board of Governors of the Federal Reserve System, *Survey of Consumer Finances* (2010).

Table 3. *Revenue of Broker-Dealers, 2010*

Revenue source	Billions of dollars
Securities commissions	\$45.6
Gains (losses) in trading and investment accounts	24.3
Profits (losses) from underwriting and selling groups	24.0
Margin interest	4.9
Revenues from sale of investment company shares	18.8
All other revenues	130.2
Total revenues	247.8
Total non-trading revenue	199.5

Source: U.S. Securities and Exchange Commission (2011).

Table 4. *Amount a Household Would Pay to Avoid High-Fee Accounts, Percent of Salary*

Avoiding	Percent of salary
Only 12b-1 fees	0.3%
Actively-managed funds	6.6
Actively-managed funds including under performance	10.0

Notes: The first row shows the percentage of salary a household investing in actively-managed funds that were not subject to 12b-1 fees would pay to avoid the imposition of those fees. The second and third rows show the percentage of salary a household investing in low-cost index funds would pay to avoid investing in funds whose returns are reduced by: 1) the excess fees on actively-managed stock and bond funds; and 2) the typical underperformance of actively-managed funds, part of which is attributable to fees and trading costs. The calculations assume constant relative risk aversion utility with a coefficient of risk aversion of five. They also assume that households face the higher fees during both the accumulation and drawdown phases.

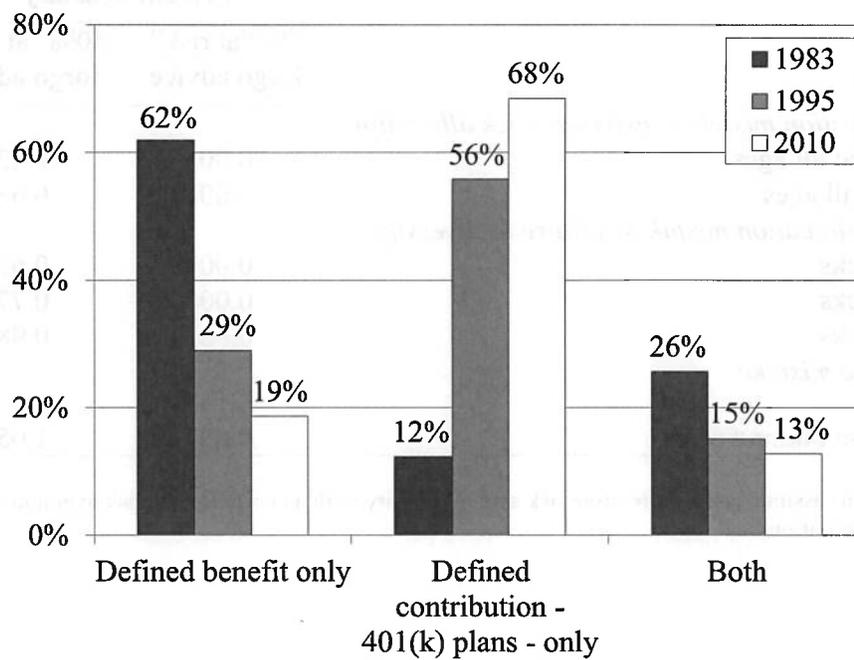
Source: Authors' calculations.

Table 5. *Amount a Household Would Pay to Avoid Mistakes, Percent of Salary*

Mistakes	Percent of salary	
	0% "at risk" forgo advice	50% "at risk" forgo advice
<i>Asset allocation mistakes: extreme stock allocation</i>		
100% at all ages	0.00	0.42
0% at all ages	0.00	0.63
<i>Portfolio allocation mistakes: failure to diversify</i>		
10 stocks	0.00	0.65
4 stocks	0.00	0.77
2 stocks	0.00	0.98
<i>Saving rate mistake</i>		
Saving nothing at all	0.00	1.05

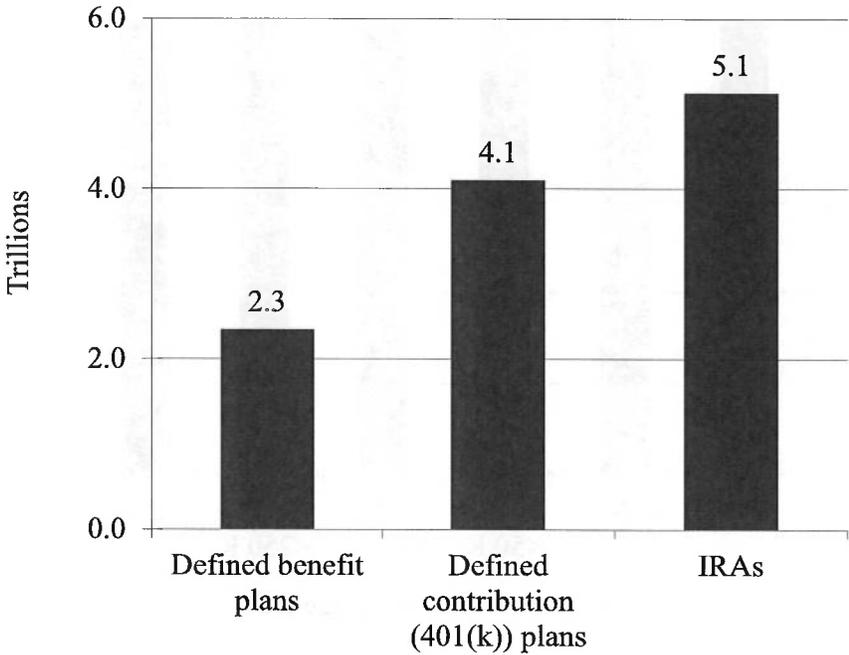
Note: The calculations assume constant relative risk aversion utility with a coefficient of risk aversion of five.
Source: Authors' calculations.

Figure 1. *Workers with Pension Coverage by Type of Plan, 1983, 1995, and 2010*



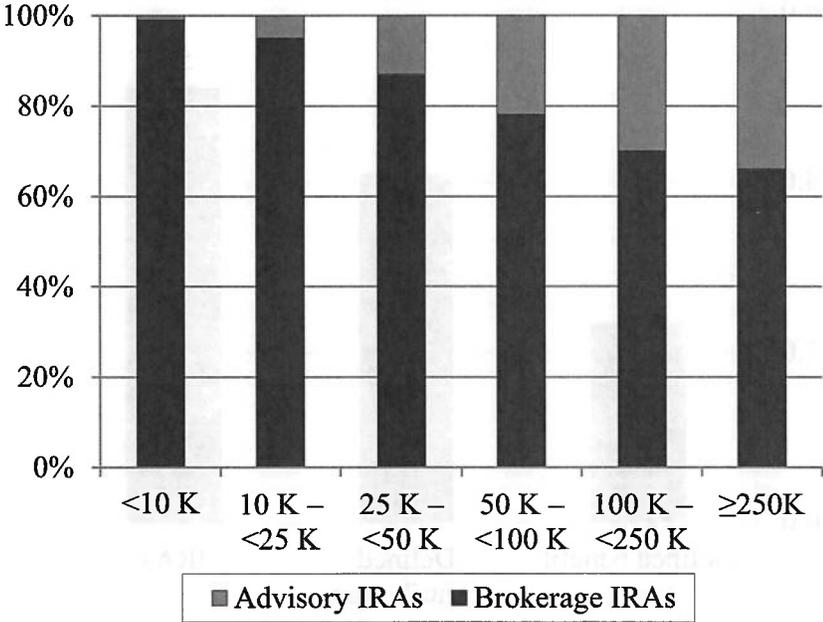
Sources: Authors' calculations based on U.S. Board of Governors of the Federal Reserve System, *Survey of Consumer Finances* (1983, 1995, and 2010).

Figure 2. *Private Retirement Assets, Trillions of Dollars, 2012 Q2*



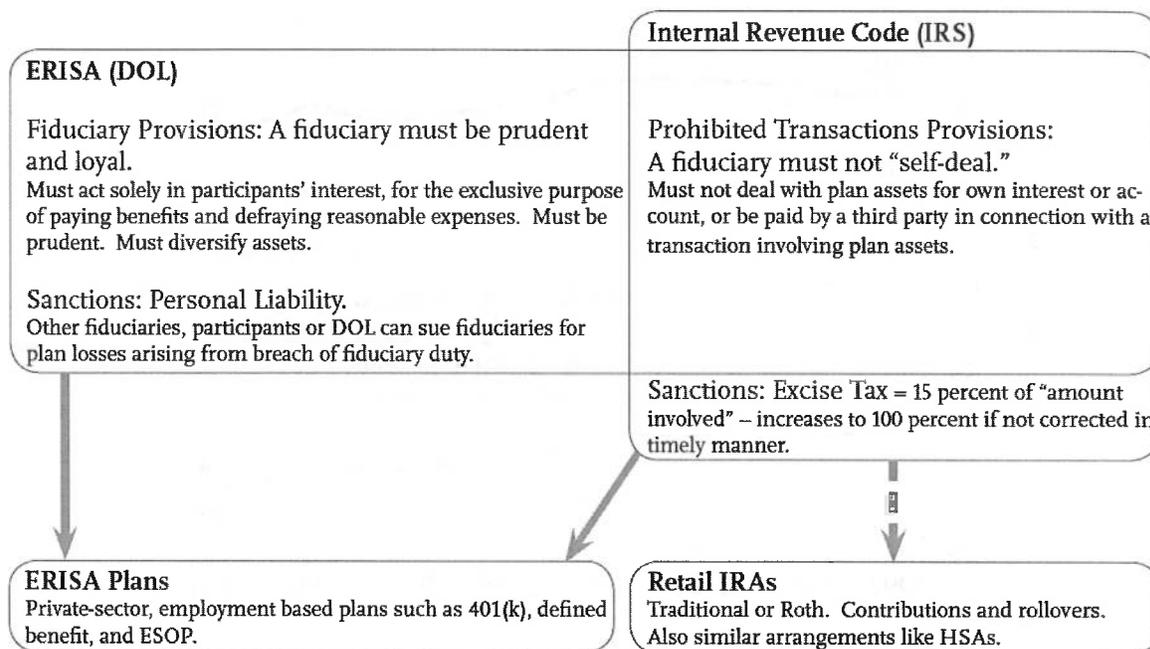
Source: U.S. Board of Governors of the Federal Reserve System, *Flow of Funds Accounts of the United States* (2012).

Figure 3. *Proportion of IRAs with Advisory and Brokerage Arrangements by Asset Size, 2010*



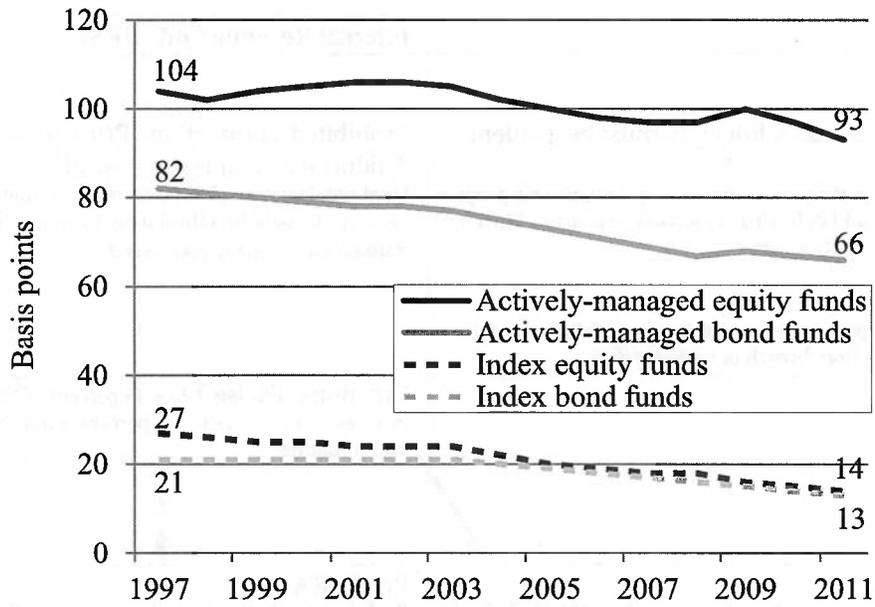
Source: Oliver Wyman (2011).

Figure 4. *Fiduciary Requirements Under ERISA and Internal Revenue Code*



Source: Authors' illustration.

Figure 5. *Expense Ratios of Actively-Managed and Index Funds, 1997-2011*



Notes: Expense ratios are measured as asset-weighted averages. The figure excludes two types of mutual funds: those available as investment choices in variable annuities and those that invest primarily in other mutual funds. Source: Investment Company Institute (2012).

Appendix

The Stochastic Dynamic Optimization Model

The optimization model assumes a married-couple household with a constant relative risk aversion (CRRA) utility function of the following form:

$$U_m(C_t^m, C_t^f) = \frac{(C_t^m + \lambda C_t^f)^{1-\gamma}}{1-\gamma}, U_f(C_t^f, C_t^m) = \frac{(C_t^f + \lambda C_t^m)^{1-\gamma}}{1-\gamma} \quad (1)$$

where λ measures the jointness of consumption; C_t^m, C_t^f denote the consumption of the husband and wife at time t ; and γ is the coefficient of risk aversion. When λ equals one, all consumption is joint. When λ equals zero, none of the household's consumption is joint. We assume that λ equals 0.5. Chiappori and Paiella (2011) argue that a CRRA utility function characterizes the risk preferences of typical households.

The household's objective is to maximize:

$$\sum_{t=22}^T \sum_{m=1}^3 \beta^{t-22} \rho_{m,t} E_t U(C_{t,m}) \quad (2)$$

where m is marital status (married, surviving male, or surviving female); β is a rate of time preference, assumed to be 0.97; $\rho_{m,t}$ is the probability of being in marital status m at time t ; and $C_{t,m}$ is consumption at time t in marital state m .⁵⁹

The household faces the following budget and non-negativity constraints:

$$W_{t+1} = (W_t + I_t + SS_t - C_t - \tau_t - FICA_t)(1 + R_{p,t+1}) \quad (3)$$

$$W_t \geq 0 \text{ for all } t \quad (4)$$

⁵⁹ We assume that the husband and wife have population average mortality for the 1970 birth cohort.

where W is wealth, I is labor income; SS is Social Security benefit; τ is federal income tax; $FICA$ is the Social Security tax; and $R_{p,t+1}$ is the portfolio return at time $t+1$. All savings are held in tax-deferred accounts. This approach reduces the number of state variables and enables us to abstract from the asset location decision. Most moderate-income households hold only small amounts of wealth outside of their 401(k)/IRA accounts. We model Social Security benefits as a non-linear function of lifetime earnings.⁶⁰ This non-linearity results in households with low lifetime earnings receiving higher rates of return on their contributions, providing partial insurance against bad labor market outcomes. This insurance may affect the optimal allocation of financial assets. We carefully model the taxation of Social Security benefits.⁶¹

The household faces uncertainty in both the labor and capital markets. We assume that households can invest in both a risk-free bond, yielding a 3-percent real interest rate, and risky stocks, with real returns log normally distributed with a mean of 6.5 percent and standard deviation of 20 percent.⁶²

We model labor income uncertainty using model and parameter values in Scholz, Seshadri, and Khitatrakun (2006). They assume that the household model of log earnings is:

⁶⁰ The Social Security retired worker benefit payable at the Full Retirement Age is a non-linear function of Average Indexed Monthly Earnings (AIME). AIME is the average of the highest 35 years' earnings, with earnings for years prior to the year in which the individual attained age 60 being indexed to the National Average Wage Index. Reflecting both the limits of computational feasibility and the focus of their research, Maurer, Mitchell, and Rogalla (2010) modeled Social Security benefits as a function of final year earnings, instead of average lifetime earnings. Their approach will likely overstate the riskiness of Social Security and potentially yield biased estimates of optimal portfolio allocations at younger ages. The magnitude of this bias is unclear ex-ante. We follow Benitez-Silva et al. (2007) and model benefits as a function of indexed lifetime earnings, making an adjustment to reflect the fact that the average of the highest 35 years earnings will be somewhat higher.

⁶¹ The taxation treatment of Social Security benefits is as follows. First, the household's "combined income" is calculated. This equals regular taxable income plus 50 percent of Social Security income. The amount of Social Security income that is taxable is the minimum of three tests: (1) 50 percent of combined income over the first threshold (\$25,000 for singles and \$32,000 for married couples), plus 35 percent of combined income over the second threshold (\$34,000 for singles and \$44,000 for married couples); (2) 50 percent of benefits plus 85 percent of combined income over the second threshold; and (3) 85 percent of benefits. See Internal Revenue Service (2012).

⁶² In reality, the real yield on long-dated bonds fluctuates considerably, while there is also considerable evidence that stock returns exhibit mean reversion. Campbell and Viceira (2002) show that long-term bonds are the true risk-free asset for long-term investors because they provide a guaranteed income. If stock returns exhibit mean reversion, then the optimal equity allocation should reflect whether expected returns are high or low relative to the long-run average. But incorporating the above features would greatly complicate our model without yielding any important additional insights.

$$\log e_j = \alpha^i + \beta_1 AGE_j + \beta_2 AGE_j^2 + u_j \quad (5)$$

$$u_j = \rho u_{j-1} + \varepsilon_j \quad (6)$$

where e_j is the observed earnings of the household i at age j in 1992 dollars; α^i is the household specific constant; AGE_j is the age of the head of the household; u_j is an AR(1) error term of the earnings equation; and ε_j is a mean zero i.i.d. normally distributed error term. They report parameter estimates in their Appendix Table A1.⁶³

We assume coefficients of risk aversion of two and five. These figures rest within the range of 2 to 10 reported in the literature, depending in part on whether the estimates are derived from portfolio theory, purchases of insurance, economic experiments, or preferences over lotteries (Chetty 2003). The retirement age is assumed to be 65.

⁶³ The model abstracts from the risk of unemployment. Incorporating the risk of unemployment would greatly complicate the analysis because we would need to both model this source of risk and incorporate social insurance programs and labor supply responses within the family.

RECENT WORKING PAPERS FROM THE
CENTER FOR RETIREMENT RESEARCH AT BOSTON COLLEGE

What Is the Long-Term Impact Of *Zebly* on Adult and Child Outcomes?

Norma B. Coe and Matthew S. Rutledge, January 2013

Sticky Ages: Why Is Age 65 Still a Retirement Peak?

Norma B. Coe, Mashfiqur Khan, and Matthew S. Rutledge, January 2013

Rethinking Optimal Wealth Accumulation and Decumulation Strategies in the Wake of the Financial Crisis

Richard W. Kopcke, Anthony Webb, and Josh Hurwitz, January 2013

Employee Mobility and Employer-Provided Retirement Plans

Gopi Shah Goda, Damon Jones, and Colleen Flaherty Manchester, November 2012

Changing Sources of Income among the Aged Population

Barry P. Bosworth and Kathleen Burke, November 2012

Holding Out or Opting Out? Deciding Between Retirement and Disability Applications in Recessions

Matthew S. Rutledge, November 2012

Automatic Enrollment, Employee Compensation, and Retirement Security

Barbara A. Butrica and Nadia S. Karamcheva, November 2012

401(k) Participant Behavior in a Volatile Economy

Barbara A. Butrica and Karen E. Smith, November 2012

Immigrant Networks and the Take-Up of Disability Programs: Evidence from U.S. Census Data

Delia Furtado and Nikolaos Theodoropoulos, October 2012

Will Delayed Retirement by the Baby Boomers Lead to Higher Unemployment Among Younger Workers?

Alicia H. Munnell and April Yanyuan Wu, October 2012

Growth in Health Consumption and Its Implications for Financing OASDI: An International Perspective

Barry P. Bosworth and Gary Burtless, September 2012

Using Participant Data to Improve 401(k) Asset Allocation

Zhenyu Li and Anthony Webb, September 2012

*All working papers are available on the Center for Retirement Research website
(<http://crr.bc.edu>) and can be requested by e-mail (crr@bc.edu) or phone (617-552-1762).*

76

TAKE IT OR LEAVE IT: UNCONSCIONABILITY OF MANDATORY PRE-DISPUTE ARBITRATION AGREEMENTS IN THE SECURITIES INDUSTRY

William Alan Nelson II, J.D., LL.M.*

INTRODUCTION	574
I. JURISPRUDENCE OF MANDATORY PRE-DISPUTE ARBITRATION	
AGREEMENTS IN THE SECURITIES INDUSTRY	576
A. <i>Enforceability in Brokerage and Advisory Contracts</i>	576
1. <i>Wilko v. Swan</i>	576
2. <i>Scherk v. Alberto-Culver</i>	578
3. <i>Shearson/American Express v. McMahon</i>	579
4. <i>Rodriguez de Quijas v. Shearson/American Express, Inc.</i>	580
B. <i>Legality of Class Action Waiver Provisions</i>	582
1. <i>AT&T Mobility, LLC v. Concepcion</i>	582
2. <i>American Express Co. v. Italian Colors Restaurant</i>	583
3. <i>FINRA v. Schwab</i>	584
II. ARGUMENTS AGAINST MANDATORY PRE-DISPUTE ARBITRATION	
AGREEMENTS.....	592
A. <i>Supreme Court Invention of a “Pro-Arbitration” Policy</i>	592
B. <i>Brokerage and Advisory Contracts are Contracts of Adhesion</i>	594
C. <i>Mandatory Pre-Dispute Arbitration Agreements Are Per Se Unconscionable</i>	595
1. <i>Involuntary Waiver of Constitutional Rights</i>	596
2. <i>Narrow Appellate Avenues Effectively Preclude Judicial Review of Arbitral Awards</i>	599

* Professorial Lecturer in Law, George Washington University Law School; Public Policy Counsel, Certified Financial Planner Board of Standards, Inc. (CFP Board). The views and opinions expressed in this Article are solely those of the author and do not necessarily represent the views and opinions of, and should not be attributed to, the CFP Board. I would like to thank the University of Pennsylvania Journal of Business Law editorial staff for their dedication in readying the Article for publication. I would also like to dedicate this Article to my beautiful wife Emily, whose constant love, support, and encouragement have sustained me beyond measure.

a. Burden of Proof Upon Aggrieved Party and Courts Extremely Deferential to Arbitrators.....	600
b. Lack of Rationale Does Not Invalidate Award.....	602
c. Impartiality, Bias of Arbitrators, and Presentation of Evidence.....	603
d. “Manifest Disregard of the Law”	604
e. Analysis.....	607
3. Unconscionability Analysis	609
III. INFORMING THE INVESTOR	612
A. <i>How a Mandatory Pre-Dispute Arbitration Agreement Would Read Under the Current Legal Framework</i>	612
IV. POSSIBLE SOLUTIONS	615
A. <i>Ending Mandatory Arbitration</i>	615
B. <i>Informed Consent and Opt-Out Provision</i>	617
CONCLUSION.....	618

INTRODUCTION

*“Counsel’s argument that the parties voluntarily agreed to arbitration and that the process saves money is also disingenuous. . . . [I]ndividuals have no recourse but to agree to an arbitration clause. Further, since the individuals seldom win and are forced to reimburse costs and attorney fees, the only ones saving money are large institutions like the claimant.”*¹

The pervasive use of mandatory pre-dispute arbitration agreements (“PDAAs”) in the securities industry is a relatively new phenomenon. Beginning in the 1980s, the Supreme Court announced that federal statutory rights, specifically claims arising under the Securities Act of 1933 (“Securities Act”) and the Securities Exchange Act of 1934 (“Exchange Act”), are subject to arbitration unless Congress has evinced a contrary intent.² An overwhelming majority of retail brokerage and many investment advisory agreements include language requiring all disputes between the customer and broker-dealer/investment adviser be resolved through arbitration³—most often with Financial Industry Regulatory

1. Wells Fargo Advisors, LLC v. Watts, 858 F. Supp. 2d 591, 595 (W.D.N.C. 2012).

2. See, e.g., Shearson/Am. Express v. McMahon, 482 U.S. 220 (1987) (holding that claims under a certain provision of the Securities Exchange Act were arbitrable under pre-dispute arbitration agreements, and that agreements to arbitrate these claims are enforceable because customers could effectively vindicate their claims against broker in arbitral forum).

3. Luis A. Aguilar, Comm’r, Sec. & Exch. Comm’n (“SEC”), Speech before the North American Securities Administrators Association (“NASAA”) Annual NASAA/SEC 19(d) Conference: Outmanned and Outgunned: Fighting on Behalf of Investors Despite

Authority (“FINRA”) Dispute Resolution. Thus, only in rare instances can an investor open a brokerage or investment advisory account without agreeing to submit to mandatory pre-dispute arbitration.

This Article begins by discussing the jurisprudence of mandatory PDAs in the securities industry, including the legality of forcing investors to waive access to class action lawsuits. The Article then transitions into a discussion of the arguments against the use of mandatory PDAs, including the Supreme Court invention of a “pro-arbitration” policy and the adhesive nature of brokerage and advisory contracts. The Article then argues that mandatory PDAs are *per se* unconscionable because they require investors to involuntarily waive certain constitutional rights and provide narrow appellate avenues that effectively preclude judicial review of arbitral awards. The Article concludes by providing an example of a mandatory PDA reflecting the current legal regime governing arbitration and possible solutions for policy-makers—the most appropriate of which is to prohibit the use of mandatory PDAs in the securities industry.

This topic is timely, especially since FINRA recently enforced a \$500,000 fine on Charles Schwab & Co. for using a mandatory PDA in contravention of FINRA rules. There has also been heightened Congressional interest in the issue, evidenced by a bill proposed by Representative Keith Ellison in the House of Representatives in 2013⁴ and a recent Senate Financial Committee report citing mandatory pre-dispute arbitration agreements as a critical issue for the SEC to investigate.⁵

The use of mandatory arbitration agreements has harmed thousands of investors over the past twenty years. A June 2007 study of more than 14,000 FINRA arbitration awards over a ten-year period (1995—2004) found that investors with significant claims suing major brokerage firms could expect to recover only *twelve percent* of the amount claimed.⁶ In a

Efforts to Weaken Investor Protections (Apr. 16, 2013), *available at* <http://www.sec.gov/News/Speech/Detail/Speech/1365171515400#.UfmMIZWdzdk> (“Currently, almost all customer agreements with brokerage firms include an arbitration clause requiring customers to arbitrate their claims in an arbitration forum—and they’re now popping-up in the investment advisory industry.”).

4. Investor Choice Act of 2013, H.R. 2998, 113th Cong. (2013).

5. STAFF OF S. COMM. ON APPROPRIATIONS, 113TH CONG., FY 2015 FINANCIAL SERVICES AND GENERAL GOVERNMENT APPROPRIATIONS BILL (Comm. Print 2014), *available at* <http://www.appropriations.senate.gov/sites/default/files/FSGG%20Report%20w%20Chart%2003REPT.PDF>.

6. See EDWARD S. O’NEAL & DANIEL R. SOLIN, SEC. LITIG. & CONSULTING GRP., MANDATORY ARBITRATION OF SECURITIES DISPUTES: A STATISTICAL ANALYSIS OF HOW CLAIMANTS FARE 15 (2007), *available at* <http://slcg.com/pdf/workingpapers/Mandatory%20Arbitration%20Study.pdf> (emphasis added) (finding a recovery percentage of just 12 percent when the claimant sought over \$250,000 against one of the top twenty brokerage firms).

2008 survey of over 3,000 individuals who participated in an arbitration (e.g., investors, advisors, attorneys), seventy-one percent of investors were dissatisfied with the outcome of the arbitration.⁷ Even with these chilling figures, it is important to note that arbitration is a proper dispute resolution process for some investors; however, individual investors should be permitted to make that choice.

I. JURISPRUDENCE OF MANDATORY PRE-DISPUTE ARBITRATION AGREEMENTS IN THE SECURITIES INDUSTRY

Even though the pervasive use of mandatory PDAAs has only arisen in the last thirty years, the issue of the legality of mandatory PDAAs stretches back to the 1950s. This Section will demonstrate the Supreme Court's evolution on the issue of mandatory PDAAs, illustrate the current view of the Supreme Court on how these agreements are enforced, and analyze recent developments concerning the use of class action waivers in mandatory PDAAs.

A. *Enforceability in Brokerage and Advisory Contracts*

1. *Wilko v. Swan*

In the Supreme Court's 1953 decision, *Wilko v. Swan*,⁸ a customer brought an action against a securities brokerage firm to recover damages under Section 12(a)(2) of the Securities Act of 1933⁹ for alleged misrepresentation in the sale of securities. The brokerage contract between

7. Jill I. Gross & Barbara Black, *When Perception Changes Reality: An Empirical Study of Investors' Views of the Fairness of Securities Arbitration*, 2008 J. DISP. RESOL. 349, 386 (2008).

8. 346 U.S. 427 (1953).

9. See Securities Act of 1933 § 12(a)(2), 15 U.S.C. § 771(a)(2) (2006 & Supp. V 2011) ("Any person who . . . sells a security (whether or not exempted by the provisions of section 77c of this title, other than paragraphs (2) and (14) of subsection (a) of said section), by the use of any means or instruments of transportation or communication in interstate commerce or of the mails, by means of a prospectus or oral communication, which includes an untrue statement of a material fact or omits to state a material fact necessary in order to make the statements, in the light of the circumstances under which they were made, not misleading (the purchaser not knowing of such untruth or omission), and who shall not sustain the burden of proof that he did not know, and in the exercise of reasonable care could not have known, of such untruth or omission, shall be liable . . . to the person purchasing such security from him, who may sue either at law or in equity in any court of competent jurisdiction, to recover the consideration paid for such security with interest thereon, less the amount of any income received thereon, upon the tender of such security, or for damages if he no longer owns the security.").

the investor and broker contained a mandatory PDAA.¹⁰ Based upon the language of the contract, the broker moved to stay the trial of the action pursuant to Section 3 of the Federal Arbitration Act (FAA)¹¹ until an arbitration, in accordance with the terms of the contract, was had.¹²

The district court held that “the agreement to arbitrate deprived [the investor] of the advantageous court remedy afforded by the Securities Act, and denied the stay.”¹³ A divided Second Circuit concluded that the Securities Act “did not prohibit the agreement to refer future controversies to arbitration, and reversed.”¹⁴ The case was subsequently taken up by the Supreme Court.

Before beginning its analysis, the Supreme Court noted that because the contract between the investor and broker evidenced a transaction in interstate commerce, there was no issue as to the applicability of the provisions of the FAA to the lawsuit.¹⁵ The Supreme Court looked at the question of “whether an agreement to arbitrate a future controversy is a ‘condition, stipulation, or provision binding any person acquiring any security to waive compliance with any provision’ of the Securities Act which § 14 declares ‘void.’”¹⁶

The Supreme Court held that Section 14 of the Securities Act voids any “‘stipulation’ waiving compliance with any ‘provision’ of the Securities Act” and that the agreement to arbitrate was “a ‘stipulation’ and . . . the right to select the judicial forum is the kind of ‘provision’ that cannot be waived under [Section] 14 of the Securities Act.”¹⁷ The Supreme Court couched this decision in terms of investor protection, specifically stating:

10. *Wilko*, 346 U.S. at 429.

11. *See* 9 U.S.C. § 3 (2012 & Supp. I 2013) (“If any suit or proceeding be brought in any of the courts of the United States upon any issue referable to arbitration under an agreement in writing for such arbitration, the court in which such suit is pending, upon being satisfied that the issue involved in such suit or proceeding is referable to arbitration under such an agreement, shall on application of one of the parties stay the trial of the action until such arbitration has been had in accordance with the terms of the agreement, providing the applicant for the stay is not in default in proceeding with such arbitration.”).

12. *Wilko*, 346 U.S. at 429.

13. *Id.* at 430.

14. *Id.*

15. *Id.*

16. *Id.* (citing Securities Act of 1933 § 14, 15 U.S.C. § 77n (2012 & Supp. I 2013)). Section 14 of the Securities Act states that “[a]ny condition, stipulation, or provision binding any person acquiring any security to waive compliance with any provision of this subchapter or of the rules and regulations of the Commission shall be void.” 15 U.S.C. § 77n.

17. *Id.* at 434–35.

While a buyer and seller of securities, under some circumstances, may deal at arm's length on equal terms, it is clear that the Securities Act was drafted with an eye to the disadvantages under which buyers labor. Issuers of and dealers in securities have better opportunities to investigate and appraise the prospective earnings and business plans affecting securities than buyers. It is therefore reasonable for Congress to put buyers of securities covered by that Act on a different basis from other purchasers.¹⁸

2. Scherk v. Alberto-Culver

The Supreme Court revisited this issue over two decades later in 1974 in *Scherk v. Alberto-Culver Co.* In *Scherk v. Alberto-Culver Co.*,¹⁹ Franz Scherk, a German businessman, sold his trademarks and interest in a European toiletries business to Alberto-Culver, Co., a U.S. corporation.²⁰ Approximately one year after the sale, Alberto-Culver tendered the property back to Scherk, because the trademarks were subject to substantial encumbrances that threatened to restrict Alberto-Culver's use of them.²¹ When Scherk refused to rescind the contract, Alberto-Culver sued Scherk in federal court in Illinois.²² Scherk subsequently filed a motion to dismiss based upon an arbitration clause in the contract of sale that provided that "any controversy or claim [that] shall arise out of this agreement or the breach thereof" would be referred to arbitration before the International Chamber of Commerce in Paris, France."²³

The U.S. District Court for the District of Illinois, relying on the Supreme Court's decision in *Wilko v. Swan*, denied Scherk's motion to dismiss and "granted a preliminary order enjoining Scherk from proceeding with arbitration."²⁴ The Seventh Circuit, similarly relying on *Wilko v. Swan*, affirmed the District Court's decision.²⁵ The Supreme Court granted certiorari to review the question of whether an arbitration agreement contained in an international sales contract was enforceable.

The Supreme Court reversed the Seventh Circuit's decision.²⁶ The Supreme Court first looked at whether the *Wilko* reasoning applied to the case at bar. The Court looked at both the Securities Act and the Securities

18. *Id.* at 435.

19. 417 U.S. 506 (1974).

20. *Id.* at 508.

21. *Id.*

22. *Id.*

23. *Id.*

24. *Id.* at 510.

25. *Id.*

26. *Id.* at 513.

Exchange Act of 1934 (Exchange Act) and found that there was no statutory counterpart for the Section 12(a)(2) action brought by the plaintiff in *Wilko*.²⁷ The Court then discussed how even if there was a statutory counterpart in the Exchange Act, the international business contract at issue involved considerations and policies significantly different than those found controlling in *Wilko*.²⁸ The Court found that “[a] contractual provision specifying in advance the forum in which disputes shall be litigated and the law to be applied is, therefore, an almost indispensable precondition to achievement of the orderliness and predictability essential to any international business transaction.”²⁹

The Supreme Court relied on the international nature of the contract at issue in *Scherk* to enforce the arbitration agreement. The Court was careful, though, to make clear that the decision “has no bearing on the scope of the substantive provisions of the federal securities laws [because] that . . . question is not presented in this case.”³⁰

3. Shearson/American Express v. McMahon

The Supreme Court revisited this issue again over a decade later in 1987 in *Shearson/American Express v. McMahon*. In *Shearson/American Express v. McMahon*,³¹ customers brought an action against a securities brokerage firm claiming that the firm violated Section 10(b) of the Exchange Act by making fraudulent and excessive trades on the customers’ accounts, making false statements, and omitting material facts from advice given to customers.³² The contract between the customers and brokerage firm contained a mandatory PDAA. Based upon the language of the contract, the brokerage firm moved to stay the trial of the action pursuant to Section 3 of the FAA until an arbitration in accordance with the terms of the contract was conducted³³

The U.S. District Court for the Southern District of New York first found that the brokerage agreement between the customers and the firm were not contracts of adhesion.³⁴ The District Court, relying on *Dean Witter Reynolds Inc. v. Byrd*³⁵ and the “‘strong national policy favoring the enforcement of arbitration agreements,’” held that the customers’ Section

27. *Id.* at 513.

28. *Id.* at 515–16.

29. *Id.* at 516.

30. *Id.* at 518 n.12.

31. 482 U.S. 220 (1987).

32. *Id.* at 223.

33. *Id.* at 223–24.

34. *Id.*

35. 470 U.S. 213 (1985).

10(b) claims were arbitrable under the terms of the agreement.³⁶ The Second Circuit, relying on *Wilko v. Swan*, reversed the District Court's decision concerning the Exchange Act claims.³⁷ The Supreme Court granted certiorari to resolve the circuit split regarding the arbitrability of Exchange Act claims.

The Supreme Court reversed the Second Circuit's decision with regard to the Exchange Act claims. The Supreme Court opened with the statement that the FAA established a "federal policy favoring arbitration" and that the Court should "rigorously enforce agreements to arbitrate."³⁸ The Court focused on its previous decision in *Wilko* and held that "[w]hile stare decisis concerns may counsel against upsetting *Wilko's* contrary conclusion under the Securities Act, we refuse to extend *Wilko's* reasoning to the Exchange Act in light of these intervening regulatory developments."³⁹ The regulatory developments cited by the Court included the competence of arbitral panels to decide securities claims and the expansive power and authority that the SEC has over the arbitral proceedings.⁴⁰ The Court concluded its decision with its interpretation of Congress's intent:

Congress did not intend for [Section] 29(a) to bar enforcement of all predispute arbitration agreements. In this case, where the SEC has sufficient statutory authority to ensure that arbitration is adequate to vindicate Exchange Act rights, enforcement does not effect a waiver of 'compliance with any provision' of the Exchange Act under [Section] 29(a).⁴¹

4. *Rodriguez de Quijas v. Shearson/American Express, Inc.*

In the Supreme Court's 1989 decision, *Rodriguez de Quijas v. Shearson/American Express, Inc.*,⁴² investors brought an action against a securities brokerage firm to recover damages under Section 12(2) of the Securities Act⁴³ for alleged misrepresentation in the sale of securities.⁴⁴

36. *Shearson*, 482 U.S. at 224.

37. *Id.*

38. *Id.* at 226 (quoting *Moses H. Cone Mem'l Hosp. v. Mercury Constr. Corp.*, 460 U.S. 1, 24 (1983) and *Dean Witter Reynolds, Inc. v. Byrd*, 470 U.S. 213, 221 (1985)).

39. *Id.* at 234.

40. 17 CFR § 240.19b-4. As part of the SEC's oversight of FINRA, FINRA must seek approval from the SEC to amend existing rules or implement new rules, including rules governing arbitration.

41. *Id.* at 238.

42. 490 U.S. 477 (1989).

43. Securities Act of 1933 § 12(2), 15 U.S.C. § 77l(2) (2006 & Supp. V 2011).

The contract between the investor and broker contained a mandatory PDAA. Based upon the language of the contract, the broker moved to stay the trial of the action pursuant to Section 3 of the Federal Arbitration Act until an arbitration in accordance with the terms of the contract was had.⁴⁵

The “District Court ordered all the claims to be submitted to arbitration except for those raised under [Section] 12(2) of the Securities Act.”⁴⁶ It held that the Securities Act claims must proceed in the court under the Supreme Court’s holding in *Wilko*.⁴⁷ The Fifth Circuit reversed, concluding that the arbitration agreement is enforceable because “[the Supreme] Court’s subsequent decisions have reduced *Wilko* to ‘obsolescence.’”⁴⁸ The Supreme Court granted certiorari to resolve the conflict among the circuit courts regarding the arbitrability of Securities Act claims.

The Supreme Court affirmed the Fifth Circuit’s decision with regard to the Securities Act claims. The Court looked at *McMahon*, which “declined to read [Section] 29(a) of the Securities Exchange Act of 1934, the language of which is in every respect the same as that in [Section] 14 of the 1933 Act . . . to prohibit enforcement of predispute agreements to arbitrate.”⁴⁹ The Court focused on its “pro-arbitration policy” by finding that “the party opposing arbitration carries the burden of showing that Congress intended in a separate statute to preclude a waiver of judicial remedies, or that such a waiver of judicial remedies inherently conflicts with the underlying purposes of that other statute.”⁵⁰ And, relying on the dissent in *Wilko*, held that “[t]here is nothing in the record before us, nor in the facts of which we can take judicial notice, to indicate that the arbitral system . . . would not afford the plaintiff the rights to which he is entitled.”⁵¹ The Court concluded by specifically holding that “*Wilko* was incorrectly decided and is inconsistent with the prevailing uniform construction of other federal statutes governing arbitration agreements in the setting of business transactions.”⁵²

The importance of the decisions in *McMahon* and *Rodriguez de Quijas* cannot be underestimated. The decisions reversed the Supreme Court’s previous decision in *Wilko* and 32 years of lower court decisions relying on *Wilko*, which had deemed Securities Act and Exchange Act

44. *Rodriguez*, 490 U.S. at 478–79.

45. *Id.* at 479.

46. *Id.*

47. *Id.*

48. *Id.*

49. *Id.* at 482.

50. *Id.* at 483.

51. *Id.* (quoting *Wilko*, 346 U.S. at 439 (Frankfurter, J., dissenting)).

52. *Id.* at 484.

claims not arbitrable. These decisions also signify the first time the Supreme Court interjected its “pro-arbitration” policy into a securities law decision. This policy will be discussed in further detail later in the Article.

B. Legality of Class Action Waiver Provisions

The Supreme Court has effectively ruled that companies may include class action waiver provisions in their brokerage and advisory contracts.⁵³ The Supreme Court’s recent allowance of class action waivers deviates from prior holdings that touted the merits of class action lawsuits.⁵⁴ Additionally, FINRA recently opposed the Supreme Court’s allowance of class action waivers and fined one of its members for using it in a brokerage contract.

1. AT&T Mobility, LLC v. Concepcion

In *Concepcion*, customers who entered into a contract with AT&T brought a class action suit alleging fraud and false advertising in connection with the carrier’s advertising practices.⁵⁵ The customer contract provided for “arbitration of all disputes between the parties, but required that claims be brought in the parties’ ‘individual capacity, and not as a plaintiff or class member in any purported class or representative proceeding.’”⁵⁶ AT&T moved to compel arbitration, citing the mandatory arbitration agreement in its customer contract that prohibited class actions.⁵⁷ The district court denied AT&T’s motion and held that the mandatory arbitration provision was unconscionable because AT&T had failed to demonstrate that the provision “adequately substituted for the deterrent effects of class actions.”⁵⁸ The Ninth Circuit affirmed, holding that the class action waiver provision was unconscionable and that the California rule was not preempted by the FAA because it was “a

53. *See, e.g.*, *AT&T Mobility LLC v. Concepcion*, 131 S. Ct. 1740 (2011) (holding that the FAA preempted a California law that would have invalidated an arbitration clause containing a class action waiver).

54. *Amchem Prods., Inc. v. Windsor*, 521 U.S. 591, 617 (1997) (quoting *Mace v. Van Ru Credit Corp.*, 109 F.3d 338, 344 (7th Cir. 1997) (“The policy at the very core of the class action mechanism is to overcome the problem that small recoveries do not provide the incentive for any individual to bring a solo action prosecuting his or her rights.”)).

55. *Concepcion*, 131 S. Ct. at 1744.

56. *Id.*

57. *Id.* at 1745.

58. *Id.* (citing *Laster v. T-Mobile USA, Inc.*, 2008 WL 5216255, at *14 (S.D. Cal. Aug. 11, 2008)).

2015]

TAKE IT OR LEAVE IT

583

refinement of the unconscionability analysis applicable to contracts generally in California.”⁵⁹

The Supreme Court, in a divided 5—4 decision, reversed the Ninth Circuit’s holding, finding that the FAA preempted state law that operated to invalidate a contractual arbitration provision. Specifically, the Supreme Court held that because the rule had been routinely applied by California courts to invalidate arbitration agreements in adhesion contracts, the rule stood as “an obstacle to the accomplishment and execution of the full purposes and objectives of Congress.”⁶⁰

2. American Express Co. v. Italian Colors Restaurant

In *Italian Colors*, the plaintiffs were merchants who accepted American Express charge cards and credit cards pursuant to contracts with American Express.⁶¹ They alleged that American Express violated the federal antitrust laws by forcing them to accept American Express credit cards at interchange rates for merchants that were approximately 30 percent higher than the rates for competing credit cards.⁶² The business contract contained “a clause that requires all disputes between the parties to be resolved by arbitration. The agreement also provide[d] that ‘[t]here shall be no right or authority for any [c]laims to be arbitrated on a class action basis.’”⁶³ The District Court granted the motion of American Express to compel individual arbitration.⁶⁴ The Second Circuit reversed and remanded for further proceedings, holding that “because respondents had established that ‘they would incur prohibitive costs if compelled to arbitrate under the class action waiver,’ the waiver was unenforceable and the arbitration could not proceed.”⁶⁵

The Supreme Court, in a 5—3 decision, held that a waiver of class arbitration was enforceable under the FAA, even if the plaintiffs’ costs of individually arbitrating their antitrust claims exceeded the potential recovery for each plaintiff.⁶⁶ The Supreme Court specifically held that “the fact that it is not worth the expense involved in *proving* a statutory remedy does not constitute the elimination of the *right to pursue* that remedy” and

59. *Id.* (quoting *Laster v. AT&T Mobility LLC*, 584 F.3d 849, 857 (9th Cir. 2009)).

60. *Id.* at 1753 (quoting *Hines v. Davidowitz*, 312 U.S. 52, 67 (1941)).

61. *Am. Express Co. v. Italian Colors Rest.*, 133 S. Ct. 2304 (2013).

62. *Id.* at 2308.

63. *Id.*

64. *Id.*

65. *Id.* (quoting *In re Am. Express Merchs. Litig.*, 554 F.3d 300, 315–16 (2d Cir. 2009)).

66. *Id.* at 2311.

that the Court's previous decision in *Concepcion* "all but resolve[d] this case."⁶⁷

3. FINRA v. Schwab

In February 2012, FINRA, which is an independent, not-for-profit organization authorized by Congress to regulate securities firms and associated persons who do business with the public, brought a disciplinary action⁶⁸ against Charles Schwab & Co., a brokerage and banking company.⁶⁹ The first cause of action alleged that Schwab, by placing a class-action waiver in its customer agreements and attempting to limit customers' ability to bring or participate in class actions, violated FINRA Rule 2268(d)(3)⁷⁰ and FINRA Code of Arbitration Procedure for Customer Disputes ("FINRA Customer Code") Rule 12204(d).⁷¹ These regulations concern a firm's ability to use a pre-dispute arbitration agreement to limit or contradict FINRA rules and enforce a class action waiver provision in its contract.

FINRA takes disciplinary action through issuance of a formal complaint. When a complaint is issued, the case is heard before a Hearing Panel that is chaired by a professional hearing officer and includes two industry representatives.⁷² A firm or individual can appeal a Hearing Panel decision to FINRA's National Adjudicatory Council (NAC), a body that is

67. *Id.* at 2312 (emphasis in original).

68. The Complaint was filed with the FINRA Office of Hearing Officers.

69. Dep't. of Enforcement v. Charles Schwab & Co., Inc., Disciplinary Proceeding No. 2011029760201 (FINRA Office of Hearing Officers Feb. 21, 2013), at 3, available at <http://www.finra.org/web/groups/industry/documents/ohodecisions/p258285.pdf>).

70. FINRA Rule 2268(d)(1) states that "[n]o predispute arbitration agreement shall include any condition that . . . limits or contradicts the rules of any self-regulatory organization . . ." FINRA RULE 2268(d)(1), FINRA MANUAL: DUTIES & CONFLICTS., FIN. INDUS. REGULATORY AUTH. (as amended in 2011), available at http://finra.complinet.com/en/display/display_main.html?rbid=2403&element_id=9955&filtered_tag=

71. FINRA Customer Code Rule 12204(d) states that:

[a] member or associated person may not enforce any arbitration agreement against a member of a certified or putative class action with respect to any claim that is the subject of the certified or putative class action until: [(1)] [t]he class certification is denied; [(2)] [t]he class is decertified; [(3)] [t]he member of the certified or putative class is excluded from the class by the court; or [(4)] [t]he member of the certified or putative class elects not to participate in the class or withdraws from the class according to conditions set by the court, if any.

FINRA RULE 12204(d), FINRA MANUAL: CODE OF ARBITRATION PROCEDURE FOR CUSTOMER DISPUTES, FIN. INDUS. REGULATORY AUTH. (as amended in 2008), available at http://finra.complinet.com/en/display/display_viewall.html?rbid=2403&element_id=4096&record_id=5174&filtered_tag=

72. *Adjudication*, FIN. INDUS. REGULATORY AUTH., available at <http://www.finra.org/Industry/Enforcement/Adjudication/> (last visited Mar. 5, 2015).

equally balanced between individuals who are in the securities business and non-industry representatives. FINRA's Board of Governors can decide whether or not to review the NAC's decision; if there is no review, the NAC decision is FINRA's final action in the matter. A firm or individual can appeal FINRA's final decision to the SEC and then to a federal court.⁷³

The Hearing Panel issued its decision in February 2013.⁷⁴ The Hearing Panel's decision concentrated primarily on two issues: (1) whether Schwab's waiver conflicts with FINRA rules, and (2), if so, whether the FAA preempts FINRA rules. The Hearing Panel found that both FINRA Rules 2268(d)(1) and (d)(3), acting in conjunction with Rule 12204 of the FINRA Customer Code, banned the use of class action waivers by FINRA members.⁷⁵ However, the Hearing Panel ultimately concluded that these Rules may not be enforced.⁷⁶ The Hearing Panel found that enforcement was foreclosed by the FAA, as construed by the Supreme Court in *Concepcion* and other decisions.⁷⁷ Because the Hearing Panel found no clear expression of congressional intent to preserve judicial class actions as an option for customer claims where there is an agreement providing for arbitration of those claims, the Hearing Panel granted Schwab's motion for summary disposition concerning the mandatory PDAA's class-action waiver provision.

Contemporaneously with the FINRA disciplinary action, Schwab filed a complaint for declaratory and preliminary and permanent injunctive relief against FINRA in the United States District Court for the Northern District of California.⁷⁸ "Schwab argued that FINRA Rule 2268(d), properly interpreted, does not prohibit class action waivers and, in the alternative, even if intended to do so, the rule's enforcement would impermissibly violate the FAA, as interpreted by the Supreme Court in *AT&T Mobility . . . and CompuCredit Corp. v. Greenwood*, 132 S. Ct. 665 (2012)."⁷⁹

FINRA filed a motion to dismiss the complaint, arguing that the federal court lacked jurisdiction. On May 11, 2012, the district court

73. *Id.*

74. Dep't. of Enforcement v. Charles Schwab & Co., Inc., Disciplinary Proceeding No. 2011029760201 (FINRA Office of Hearing Officers Feb. 21, 2013), at 1, available at <http://www.finra.org/web/groups/industry/documents/ohodecisions/p258285.pdf>.

75. *Id.* at 23–24.

76. *Id.* at 40–41.

77. *Id.*

78. *Charles Schwab & Co v. Fin. Indus. Regulatory Auth.*, 861 F. Supp. 2d 1063 (N.D. Cal. 2012).

79. Dep't of Enforcement v. Schwab, Disciplinary Proceeding No. 2011029760201 (FINRA Board of Governors Apr. 24, 2014), at 5 n.8, available at <http://www.finra.org/web/groups/industry/@ip/@enf/@ad/documents/industry/p493598.pdf> [hereinafter Board of Governors Decision]; *Charles Schwab & Co*, 861 F. Supp. 2d at 1067.

granted FINRA's motion and dismissed Schwab's complaint. The court determined that it lacked subject matter jurisdiction because Schwab did not exhaust the administrative remedies established by the Exchange Act and that Schwab failed to show it was entitled to an exception from the general exhaustion requirement.⁸⁰

This is an important procedural issue, because for any companies who want to challenge this type of disciplinary action, generally they must go through the administrative process before being able to utilize the judicial system. Many companies might find the burden cost inefficient and may just remove these provisions from their mandatory PDAAs.

The Board of Governors found that the two main issues regarding the enforceability of Schwab's pre-dispute arbitration agreements were: (1) whether FINRA rules preserve the right for investors to participate in judicial class actions and (2) whether the FAA applies to FINRA arbitration rules and preempts enforcement of those rules.⁸¹

The Board of Governors addressed these two issues in its decision and ultimately overruled the Hearing Panel decision, stating that:

We uphold these FINRA rules and find that Schwab's inclusion of a mandatory waiver of participation in judicial class actions, as well as its restriction of an arbitrator's power to join together individual claims violates NASD and FINRA rules. Because we determine that the FAA does not preclude FINRA's enforcement of its rules, we reverse the Hearing Panel's dismissal of the first two causes of action. We remand this matter to the Hearing Panel to determine appropriate sanctions.⁸²

In arriving at this decision, the Board of Governors first looked at the language of the rules in question (FINRA Rule 2268 and FINRA Customer Code Rule 12204). FINRA Rule 2268 sets the requirements for FINRA members when using pre-dispute arbitration agreements. Rule 2268(f) requires all pre-dispute arbitration agreements for customer accounts to state that no person may bring a class action in arbitration, nor seek to enforce a pre-dispute arbitration agreement against a person who has initiated a judicial class action or is a member of a putative class until class certification issues are decided.⁸³ The decision also focused on Rule

80. Board of Governors Decision, *supra* note 79, at 5 n.8; *Charles Schwab & Co.*, 861 F. Supp. 2d at 1069.

81. Board of Governors Decision, *supra* note 79, at 3.

82. *Id.*

83. FINRA RULE 2268(f), FINRA MANUAL: DUTIES & CONFLICTS, FIN. INDUS. REGULATORY AUTH. (as amended in 2011), available at <http://finra.complanet.com>

2268(d)(3) which prohibits members from incorporating conditions in a pre-dispute arbitration agreement that “limit[] the ability of a party to file any claim in court permitted to be filed in court under the rules of the forums in which a claim may be filed under the agreement.”⁸⁴

Schwab argued that FINRA Rule 2268(d)(3) cannot be referring to class actions when the rule language uses the term “claim” because class actions are procedural mechanisms and not claims.⁸⁵ To determine the meaning of the word “claim”, the Board of Governors turned to the definition set out in FINRA Customer Code Rule 12100(d), which defines a “claim” as “an allegation or request for relief.”⁸⁶ Rule 12204 of the Customer Code is titled “Class Action Claims” and specifically addresses the status of class action claims in FINRA arbitration, stating that “[c]lass action claims may not be arbitrated under the [Customer] Code.”⁸⁷

The Board of Governors then proceeded to examine the rule-making history. It specifically focused on the intent behind Rule 12204 and found that the SEC stated “in all cases, class actions are better handled by the courts and that investors should have access to the courts to resolve class actions efficiently” and “[w]ithout access to class actions in appropriate cases, both investors and broker-dealers have been put to the expense of wasteful, duplicative litigation.”⁸⁸

The Board of Governors ultimately determined that “a securities-law claim brought as a class action therefore is a category of claim that was intended to be filed in court under FINRA rules.”⁸⁹ The Board of Governors further determined that “Rule 12204 of the Customer Code was intended to preserve investor access to the courts to bring or participate in judicial class actions, and that through its Waiver, Schwab violated FINRA Rules 2268(d)(1) and (d)(3), and Rule 12204 of the Customer Code.”⁹⁰

Even though the Board of Governors found that there was a breach of its rules, they still had to determine whether the FAA pre-empted enforcement of those rules. As a threshold matter, the Board of Governors found that the FAA applied to the case because it governs virtually every

/en/display/display_main.html?rbid=2403&element_id=9955&filtered_tag=

84. FINRA RULE 2268(d)(3), FINRA MANUAL: DUTIES & CONFLICTS, FIN. INDUS. REGULATORY AUTH. (as amended in 2011), available at http://finra.complinet.com/en/display/display_main.html?rbid=2403&element_id=9955&filtered_tag=

85. Board of Governors Decision, *supra* note 79, at 8.

86. *Id.* at 10.

87. *Id.*

88. *Id.* at 14 (quoting Order Approving Proposed Rule Change Relating to the Exclusion of Class Actions from Arbitration Proceedings, Exchange Act Release No. 31371, 1992 WL 324491, at *3).

89. *Id.* at 10.

90. *Id.* at 15.

arbitration agreement arising out of a commercial transaction, and Schwab's customer transactions were no exception.⁹¹ This decision first makes clear that even though there is a long-standing policy of favoring arbitration agreements, the FAA has limits. The Board of Governors cited *CompuCredit Corp. v. Greenwood*, where "the Supreme Court reaffirmed the principle that the mandate of the FAA is not absolute, explaining that it may be 'overridden by a contrary congressional command.'"⁹²

To meet the standard under *Greenwood*, the decision specifically found that "both the text of the Exchange Act and the rulemaking history of NASD's proposal to adopt what are currently FINRA Rule 2268 and Rule 12204 of the Customer Code demonstrate a statutorily authorized intent to overcome the FAA."⁹³

The Board of Governors looked at the Hearing Panel's determination that found no congressional intent to preserve judicial class actions as an option for customer claims under FINRA rules. They reversed the Hearing Panel and found that FINRA's authority to promulgate rules is a Congressional command.⁹⁴ The decision cited to *CompuCredit*, where the Supreme Court cited the Consumer Financial Protection Bureau's ("CFPB") authority to regulate predispute arbitration agreements as an example of a contrary congressional command. The Supreme Court found that "Congress both delegated to an agency and allowed the agency to exercise judgment. Consequently, we conclude that Congress's granting of authority to the SEC to approve of SRO limitations on arbitration agreements is equally as valid as its granting of authority to the CFPB."⁹⁵

The Board of Governors then cited the Ninth Circuit's decision in *Credit Suisse First Boston Corp. v. Grunwald* for the proposition that "FINRA rules have the force and effect of a federal regulation for the

91. *Id.* at 16. They specifically noted that the Supreme Court had applied the FAA to securities arbitrations under the Exchange Act. *Id.*; see *Shearson/Am. Express, Inc. v. McMahon*, 482 U.S. 220 (1987) (indicating courts should favor enforcement of arbitration agreements).

92. Board of Governors Decision, *supra* note 79, at 18 (quoting *CompuCredit Corp. v. Greenwood*, 132 S. Ct. 665, 669 (2012)).

93. *Id.*

94. *Id.* at 20.

95. *Id.* at 20. "The Exchange Act manifestly gives FINRA the ability to propose, and the SEC the authority to approve, rules that govern which claims will be submitted to arbitration and which will not. The rulemaking history of FINRA's rules evidences that the SEC was approving FINRA's restrictions on predispute arbitration agreements pursuant to the Exchange Act. The SEC followed Congress's designated process for a FINRA rule to be approved. The Exchange Act's broad authorization encompassing FINRA arbitration rules that are approved by the SEC constitutes the Supreme Court's required congressional command to overcome the general mandate of the FAA to enforce arbitration agreements." *Id.* at 21.

purposes of resolving federal conflicts of law.”⁹⁶ In *Grunwald*, the Ninth Circuit held that NASD’s arbitration rules regarding disclosure and disqualification of arbitrators, which had been approved by the SEC, preempted conflicting state law requirements regarding arbitrators.⁹⁷ The Board of Governors surveyed other cases where conflicts existed between the antitrust laws and the securities laws, including SEC and FINRA rules, and found that securities laws prevailed when the Court determined that the antitrust laws would produce conflicting guidance in an area that is addressed by the securities laws.⁹⁸

The Board of Governors reasoned that:

[t]he Supreme Court’s holdings that federal antitrust laws are impliedly repealed when they conflict with securities laws support the conclusion that the Exchange Act, effectuated through FINRA rules, overrides the FAA here. First, the SEC, through its oversight of the FINRA arbitration forum and its review of FINRA rules, has consistently exercised its authority to oversee the arbitration process. Second, the Schwab dispute represents a direct conflict between the FAA, which mandates enforcement of arbitration agreements according to their terms, and FINRA rules, which require broker dealers to preserve judicial class actions for investors. Third, FINRA’s arbitration rules for customer disputes with FINRA firms and associated persons address a critical aspect of investor protection under the Exchange Act, namely in what forum a customer class action will be litigated. In reconciling the conflict between FINRA arbitration rules that prohibit use of a predispute arbitration agreement to eliminate judicial class actions and the FAA’s enforcement of class action waivers, we find—based on the SEC’s approval orders—that FINRA’s rules are in furtherance of the Exchange Act’s protection of investors. This core aspect of the Exchange Act prevails over the FAA.⁹⁹

Schwab sought to use a series of recent Supreme Court opinions involving pre-dispute arbitration agreements between customers and

96. *Id.* at 21 (citing *Credit Suisse First Boston Corp. v. Grunwald*, 400 F.3d 1119, 1132 (9th Cir. 2005)).

97. *Id.* (citing *Grunwald*, 400 F.3d at 1132).

98. *Id.* (citing *Gordon v. N.Y. Stock Exch., Inc.*, 422 U.S. 659 (1975); *United States v. NASD, Inc.*, 422 U.S. 694 (1975)).

99. *Id.* at 22.

corporations to invalidate two of FINRA's longstanding arbitration rules.¹⁰⁰ However, after careful consideration, the Board of Governors found that FINRA's rules had not been invalidated by recent Supreme Court holdings related to pre-dispute arbitration agreements.¹⁰¹ In so doing, the Board of Governors first looked at *Concepcion*, which involved a state law rule that invalidated class action waivers when one party had superior bargaining power and was alleged to have caused small damages to many customers.¹⁰² FINRA held that "the line of cases in which courts have invalidated state laws based on federal preemption, however, do not apply to this case. FINRA arbitration rules are not creatures of state law nor should they be treated as equivalent to state law."¹⁰³

The Board of Governors then looked at *American Express Co. v. Italian Colors Restaurant*, which reaffirmed that "the FAA 'requires the enforcement of arbitration agreements containing class action waivers' unless a law contains a contrary congressional command."¹⁰⁴ FINRA distinguished *Italian Colors* because, in that case, the Supreme Court found that "[p]re-dispute arbitration waivers that have the effect of making an individual party bear the expense of proving a violation . . . are valid and enforceable."¹⁰⁵ However, the Board of Governors found that *Italian Colors* did not apply to Schwab because there is no contention that customers must be allowed to pursue judicial class actions as a matter of efficiency.¹⁰⁶

Schwab also relied on *McMahon*, which, as noted above, is the seminal case applying the FAA to securities arbitrations. Schwab contended that the holding in *McMahon*, that arbitration agreements were not invalid waivers of Exchange Act claims (those claims could be pursued in arbitration), mandates that class action waivers are valid.¹⁰⁷ But the Board of Governors found that the anti-waiver provision under the Exchange Act at issue in *McMahon* only prohibited a waiver of the substantive obligations imposed by the Exchange Act. The Board of Governors reasoned that "Rule 12204 of the Customer Code, in contrast, preserves investor access to the courts to bring or participate in judicial class actions. Schwab's class action waiver directly violates FINRA rules.

100. *Id.* at 23-24.

101. *Id.* at 23.

102. *See supra* Section II.B.1 (discussing the *AT&T Mobility* decision).

103. Board of Governors Decision, *supra* note 79, at 23.

104. *Id.* (quoting *Am. Express Co. v. Italian Colors Rest.*, 133 S. Ct. 2304 (2013)).

105. *Id.* (citing *Italian Colors*, 133 S. Ct. at 2311).

106. *Id.* at 23-24.

107. *Id.* at 24.

It is not invalid because it requires a customer to waive a substantive obligation imposed by the Exchange Act.”¹⁰⁸

This decision has the ability to be a game-changer in the mandatory arbitration discussion. FINRA is asserting its authority to invalidate class action waivers in the mandatory PDAAs of its members. The fact that Schwab paid a \$500,000 fine and waived its right to appeal the decision is very telling.¹⁰⁹ This large fine may deter other companies from placing class action waivers in their pre-dispute arbitration agreements.

It is important to note, however, that this is a decision made by FINRA, not a court of law. If a similar case rises to the courts, it will be much harder for plaintiffs to argue against enforcement of the class action waiver, due to the negative precedent set by *Concepcion*, *Italian Colors*, and the Supreme Court’s long-standing policy favoring arbitration. This notion is reinforced by the recent decision in *Torres v. United Healthcare Servs., Inc.*, where a New York District Court found that the National Labor Relations Board’s (“NLRB”) interpretation of the FAA was outside its expertise and not entitled to deference.¹¹⁰ Courts have also found that they are “not obligated to defer to an agency’s interpretation of Supreme Court precedent under *Chevron* or any other principle.”¹¹¹

It is also important to note that this action was mainly brought about by customer complaints. Schwab admitted in news articles that:

[o]ver the last year, we heard clearly that a number of our clients and members of the general public have strong feelings about maintaining access to class-action lawsuits. In a business like ours where our reputation and public trust are key to our success, we take perspectives like those very seriously.¹¹²

This Article contends that the reputational risk faced by companies will be more important than any judicial decision. This case helped confirm that principle. For example, even though both the FINRA Hearing Panel and courts allowed the class action waiver to be included in the pre-dispute arbitration agreement, investor complaints and bad publicity

108. *Id.* (quoting *Shearson/Am. Express, Inc. v. McMahon*, 482 U.S. 220, 228 (1987)).

109. *See id.* at 28. (detailing the sanctions imposed by FINRA’s Board of Governors on Schwab).

110. *Torres v. United Healthcare Servs., Inc.*, 920 F. Supp. 2d 368, 378 (E.D.N.Y. 2013).

111. *N.Y. N.Y., LLC v. N.L.R.B.*, 313 F.3d 585, 590 (D.C. Cir. 2002).

112. *Jed Horowitz & Suzanne Barlyn, Schwab Drops Ban on Clients Filing Class-Action Lawsuits*, REUTERS, (Apr. 24, 2014), <http://www.reuters.com/article/2014/04/24/schwab-classaction-settlement-idUSL2N0NG1XJ20140424>.

ultimately caused Schwab to remove the provision, even if for a short period of time.

II. ARGUMENTS AGAINST MANDATORY PRE-DISPUTE ARBITRATION AGREEMENTS

This Article presents multiple arguments against the imposition of mandatory PDAAs in brokerage and advisory contracts including: (1) the Supreme Court's invention of a "pro-arbitration" policy; (2) involuntary relinquishment of certain Constitutional rights due to mandatory PDAAs; and (3) the lack of options available to investors to overturn an erroneous arbitration award effectively preclude judicial review.

A. Supreme Court Invention of a "Pro-Arbitration" Policy

As discussed in the cases in Section I, over the past thirty years the Supreme Court has developed the idea of a "pro-arbitration" policy.¹¹³ Over the past three decades, the practice of using arbitration in lieu of litigation has not only garnered approval from the Supreme Court, but has been encouraged by the Court.¹¹⁴ It is important to note the distinction between voluntary and mandatory arbitration. "While arbitration has been used as a dispute resolution technique for thousands of years, in the past it has been agreed to knowingly and voluntarily, typically by two or more businesses. The involuntary imposition of arbitration in lieu of open court procedures is a new and most controversial phenomenon."¹¹⁵

In its 1982 decision in *Moses H. Cone*, the Supreme Court stated that the FAA was a "congressional declaration of a liberal federal policy favoring arbitration agreements, notwithstanding any state substantive or procedural policies to the contrary."¹¹⁶ This was the first instance in which the Supreme Court enunciated a federal policy that favors arbitration of commercial disputes. This statement by the Court, and subsequent holdings in *McMahon* and *Rodriguez de Quijas*, have provided businesses the opportunity to compel arbitration in contexts where they previously thought arbitration agreements would not be enforced.

113. See *supra* Section I (illustrating judicial opinions in support of a pro-arbitration paradigm).

114. See, e.g., *Moses H. Cone Mem'l Hosp. v. Mercury Constr. Corp.*, 460 U.S. 1, 24 (1982) (noting that the FAA established a "federal policy favoring arbitration" and that the courts should "rigorously enforce agreements to arbitrate.").

115. Jean R. Sternlight, *Creeping Mandatory Arbitration: Is It Just?*, 57 STAN. L. REV. 1631, 1631–32 (2005) (citations omitted).

116. *Moses H. Cone Mem'l Hosp. v. Mercury Constr. Corp.*, 460 U.S. 1, 24 (1982).

This Article contends that the FAA that Congress adopted over eighty years ago bears little resemblance to the Supreme Court's current interpretation of the FAA. The legislative history of the FAA indicates that the intention behind the law was to allow courts to enforce arbitration agreements between merchants.¹¹⁷ Interestingly, members of Congress voiced concerns that the FAA would be used against unsuspecting investors. For example, Montana Senator Thomas Walsh, during a Joint Committee hearing, voiced his concern that:

a great many of these contracts that are entered into are really not voluntary things at all A man says, 'These are our terms. All right, take it or leave it.' Well, there is nothing for the man to do except to sign it; and then he surrenders his right to have his case tried by the court, and has to have it tried before a tribunal in which he has no confidence at all."¹¹⁸

The Supreme Court seemed to overlook the legislative history that cautioned against enforcement of mandatory PDAs on parties with unequal bargaining power.

The FAA's primary purpose was to ensure enforcement of arbitration agreements, not to encourage businesses to place mandatory PDAs in their contracts with investors.¹¹⁹ The legislative history and case law reflect that "arbitration is a matter of contract" and that the FAA puts arbitration agreements "on equal footing with other contracts"¹²⁰ and "creates no new legislation, grants no new rights, except a remedy to enforce an agreement in commercial contracts and in admiralty contracts."¹²¹

It is also important to note that though the Supreme Court speaks of a pro-arbitration policy in all of the cited cases,¹²² the Court never states that Congress intended a pro-*mandatory* arbitration policy. Based upon the

117. *See Arbitration of Interstate Commercial Disputes: Hearing of S. 1005 and H.R. 646 Before the J. Comm. of Subcomms. on the Judiciary*, 68th Cong. 16 (1924) (statement of Julius Cohen, Gen. Counsel, N.Y. State Chamber of Commerce) (discussing the applicability of such agreements to federal court decisions that bind domestic and international trade agreements).

118. *Sales and Contracts to Sell in Interstate and Foreign Commerce, and Federal Commercial Arbitration: Hearing on S. 4213 and S. 4214 before a Subcomm. of the S. Comm. on the Judiciary*, 67th Cong., 4th Sess., 9 (1923) (statement of Sen. Thomas Walsh).

119. *Dean Witter Reynolds Inc. v. Byrd*, 470 U.S. 213, 219–20 (1985) (concluding that the FAA was designed "to overrule the judiciary's long-standing refusal to enforce agreements to arbitrate").

120. *Janiga v. Questar Capital Corp.*, 615 F.3d 735, 740, 742 (7th Cir. 2010).

121. 65 CONG. REC. 1931 (1924) (statement of Rep. George S. Graham) (regarding the enforcement of arbitration agreements).

122. *See supra* Section I.A.

case law and the legislative history of the FAA, this is an important distinction. The Supreme Court has consistently held that arbitration is a matter of *consent*. In *Volt Information Sciences, Inc. v. Board of Trustees*, the Supreme Court held that under the FAA, arbitration “is a matter of consent, not coercion, and parties are generally free to structure their arbitration agreements as they see fit.”¹²³ In *Stolt-Nielsen S.A. v. AnimalFeeds International Corp.*, the Supreme Court reaffirmed this idea by stating clearly that “[t]he Court ties the requirement of affirmative authorization to ‘the basic precept that arbitration is a matter of consent, not coercion.’”¹²⁴ As noted in the next Section, brokerage and advisory contracts are offered on a “take it or leave it” basis and do not require the type of consent required by the Supreme Court.

B. Brokerage and Advisory Contracts are Contracts of Adhesion

A contract of adhesion is defined as a “standard-form contract prepared by one party, to be signed by another party in a weaker position . . . who adheres to the contract with little choice about the terms.”¹²⁵ “An adhesion contract is typically a standardized form ‘offered to consumers of goods and services on essentially a ‘take it or leave it’ basis without affording the consumer a realistic opportunity to bargain and under such conditions that the consumer cannot obtain the desired product or services except by acquiescing in the form contract.’”¹²⁶ Courts look at multiple factors to determine whether a contract is one of adhesion, including whether there was great disparity in bargaining power, no opportunity for negotiation, or services that could not be obtained elsewhere.¹²⁷

Courts have found that brokerage and advisory contracts are generally contracts of adhesion.¹²⁸ “Adhesion contracts generally warrant heightened judicial scrutiny because the drafting party is in a superior bargaining

123. *Volt Info. Scis., Inc. v. Bd. of Trs.*, 489 U.S. 468, 479 (1989).

124. *Stolt-Nielsen S.A. v. AnimalFeeds Int’l Corp.*, 559 U.S. 662, 697–98 (2010) (quoting *Volt*, 489 U.S. at 479).

125. *Definition of Adhesion Contract*, BLACK’S LAW DICTIONARY (9th ed. 2009), available at Westlaw BLACKS.

126. *Broemmer v. Abortion Servs. of Phx., Ltd.*, 840 P.2d 1013, 1015 (1992) (quoting *Wheeler v. St. Joseph Hosp.*, 63 Cal. App. 3d 345, 356 (1976)); see *Siegelman v. Cunard White Star Ltd.*, 221 F.2d 189, 204 (2d Cir. 1955) (noting that with one-sided control of a contract’s terms “the usual contract rules, based on the idea of ‘freedom of contract,’ cannot be applied rationally”).

127. *Brenner v. Nat’l Outdoor Leadership Sch.*, 20 F. Supp. 3d 709, 717 (D. Minn. 2014).

128. See, e.g., *Shearson/Am. Express, Inc. v. McMahon*, 482 U.S. 220, 230 (1987) (“arbitration clauses in securities sales agreements generally are not freely negotiated”).

position,¹²⁹ and although they will not be found unconscionable in every case, “an adhesion contract is procedurally unconscionable and unenforceable ‘when the terms are patently unfair to the weaker party.’”¹³⁰ The Supreme Court, however, has stated that “[m]ere inequality in bargaining power . . . is not a sufficient reason to hold that arbitration agreements are never enforceable.”¹³¹ Enforceability also “depends upon whether the terms of the contract are beyond the reasonable expectations of an ordinary person”¹³²

The preceding Sections reflect the contradictory nature of the Supreme Court’s evolution on mandatory PDAAs. The Court has gone to great lengths to show that arbitration is a matter of consent, but then subsequently allows mandatory PDAAs in adhesion contracts where the investor may not even be aware of the clause binding them to future arbitration. Generally, an individual cannot consent to a contract provision that he or she is not aware of¹³³ or, in the alternative, to which he or she has not agreed.¹³⁴ The Supreme Court has also stated that “the FAA’s pro-arbitration policy does not operate without regard to the wishes of the contracting parties,”¹³⁵ which reflects an implicit understanding that the FAA’s presumption in favor of arbitration should only apply when both parties explicitly and voluntarily state their intention to enter into arbitration.

C. Mandatory Pre-Dispute Arbitration Agreements Are Per Se Unconscionable

Due to the pervasive use of adhesion contracts in the securities industry, investors are effectively forced to sign an agreement that binds

129. *Rivera v. Am. Gen. Fin. Servs., Inc.*, 259 P.3d 803, 817 (N.M. 2011).

130. *Id.* (citation omitted).

131. *Gilmer v. Interstate/Johnson Lane Corp.*, 500 U.S. 20, 33 (1991).

132. *Howell v. NHC Healthcare–Fort Sanders, Inc.*, 109 S.W.3d 731, 734 (Tenn. Ct. App. 2003).

133. *O’Connor v. Combined Ins. Co. of Am.*, 441 Fed. App’x. 362, 366 (6th Cir. 2011) (quoting *In re Brown*, 287 B.R. 676, 682 (E.D. Mich. 2001)) (“It is axiomatic that a party cannot consent to terms of which it is unaware.”).

134. The Third Circuit’s general rule is that a party need not arbitrate unless it has agreed to such arbitration. *Fencourt Reinsurance Co., Ltd. v. ITT Indus., Inc.*, 2008 WL 2502139, at *11 (E.D. Pa. 2008) (quoting *Bel-Ray Co., Inc. v. Chemrite (Pty) Ltd.*, 181 F.3d 435, 444 (3d Cir. 1999) (“Arbitration is strictly a matter of contract. If a party has not agreed to arbitrate, the courts have no authority to mandate that he do so.”); *cf. Adam v. Merrill Lynch Pierce Fenner & Smith*, 888 F.2d 696, 701 (10th Cir. 1989) (holding that a broker’s failure to verbally disclose arbitration provisions did not render the provisions unenforceable because “the law presumes that one has read that which he has signed.”).

135. *Mastrobuono v. Shearson Lehman Hutton, Inc.*, 514 U.S. 52, 57 (1995).

them to all the terms contained therein, including mandatory PDAAs.¹³⁶ Mandatory PDAAs are not within investors' reasonable expectations and should be deemed *per se* unconscionable because they require investors' involuntary waiver of certain constitutional rights and provide narrow appellate avenues that effectively preclude judicial review of arbitral awards.

1. Involuntary Waiver of Constitutional Rights

This Article contends that by forcing unwitting investors into arbitration before a dispute arises, they are being forced to relinquish certain constitutional rights. As noted throughout this Article, arbitration is a perfectly legal and efficient method to resolve disputes; however, investors should not be deprived of that choice. Individuals may prospectively waive their constitutional rights to access the court system and due process of law; this principle is not at issue. The issue that arises with mandatory PDAAs is the manner in which those rights are relinquished by the investor.

The Seventh Amendment provides that “[i]n suits at common law, where the value in controversy shall exceed twenty dollars, the right of trial by jury shall be preserved”¹³⁷ This provision governs proceedings in federal court and not state court, but each state has a provision in its constitution that is analogous.¹³⁸ An individual may waive his or her constitutional right to a jury trial as long as that waiver is “knowing and voluntary.” Under the knowing and voluntary standard, “for a waiver to be effective it must be clearly established that there was ‘an intentional relinquishment or abandonment of a known right or privilege.’”¹³⁹ Courts will generally look to:

- (1) the conspicuousness of the [waiver] provision in the contract;
- (2) the level of sophistication and experience of the parties entering into the contract;
- (3) the opportunity to negotiate terms

136. N. AM. SEC. ADM'RS ASS'N (“NASAA”), MANDATORY BINDING ARBITRATION: IS IT FAIR AND VOLUNTARY? (2009), available at <http://www.nasaa.org/wp-content/uploads/2011/07/NASAA-Arbitration-Statement-9.15.09.pdf> (“Today, almost every broker-dealer includes in their customer agreements, a predispute arbitration provision that forces public investors to submit all disputes that they may have with the firm and/or its associated persons to mandatory arbitration.”).

137. U.S. CONST. amend. VII.

138. See, e.g., ALA. CONST. of 1901, art. I, § XI (“the right of trial by jury shall remain inviolate.”).

139. *Brookhart v. Janis*, 384 U.S. 1, 4 (1966) (quoting *Johnson v. Zerbst*, 304 U.S. 458, 464 (1938)).

of the contract; (4) the relative bargaining power of each party; and (5) whether the waiving party was represented by counsel.¹⁴⁰

The Supreme Court has traditionally been wary of waivers of constitutional rights in contracts of adhesion. In *Fuentes v. Shevin*, the Supreme Court found that there was “no showing whatever that the appellants were actually aware or made aware of the significance of the fine print now relied upon as a waiver of constitutional rights” and that “a waiver of constitutional rights in any context must, at the very *least*, be clear.”¹⁴¹ Similarly, in *Teamsters v. Terry*, the Supreme Court held that “[m]aintenance of the jury as a fact-finding body is of such importance and occupies so firm a place in our history and jurisprudence that any seeming curtailment of the right to a jury trial should be scrutinized with the utmost care.”¹⁴²

Even with the Supreme Court’s wariness on the issue of waivers of constitutional rights in contracts of adhesion, courts have held that an individual’s waiver of his or her constitutional right to a jury trial under a mandatory PDAA does not have to meet the traditional “knowing and voluntary” standard for waiver. In *Caley v. Gulfstream*, the Eleventh Circuit held that “general contract principles govern the enforceability of arbitration agreements and that no heightened ‘knowing and voluntary’ standard applies, even where the covered claims include federal statutory claims generally involving a jury trial right.”¹⁴³ Courts have also found that when a party enters into a valid agreement to arbitrate, the party is not entitled to a jury trial or to a judicial forum for covered disputes. In *Mitsubishi Motors Corp.*, the Supreme Court held that “[b]y agreeing to arbitrate a statutory claim, a party does not forego the substantive rights afforded by the statute; it only submits to their resolution in an arbitral, rather than a judicial, forum.”¹⁴⁴ Likewise in *American Heritage Life Ins. Co.*, the Fifth Circuit held that the Seventh Amendment does not “confer the right to a trial, but only the right to have a jury hear the case once it is

140. *Allyn v. W. United Life Assurance Co.*, 347 F. Supp. 2d 1246, 1252 (M.D. Fla. 2004) (citations omitted).

141. *Fuentes v. Shevin*, 407 U.S. 67, 95 (1972) (emphasis in original).

142. *Chauffeurs, Teamsters & Helpers, Local No. 391 v. Terry*, 494 U.S. 558, 565 (1990) (quoting *Dimick v. Schiedt*, 293 U.S. 474, 486 (1935)).

143. *Caley v. Gulfstream Aerospace Corp.*, 428 F.3d 1359, 1372 (11th Cir. 2005) (citing *Am. Heritage Life Ins. Co. v. Orr*, 294 F.3d 702, 711 (5th Cir. 2002)) (“holding that parties’ arbitration agreement, which stated that the right to a trial and to a jury were waived, validly waived those trial rights, and rejecting arguments that a heightened ‘clear and unmistakable’ or ‘voluntary, knowing, and intelligent’ standard applied to the waiver”).

144. *Mitsubishi Motors Corp. v. Soler Chrysler-Plymouth, Inc.*, 473 U.S. 614, 628 (1985).

determined that the litigation should proceed before a court. If the claims are properly before an arbitral forum pursuant to an arbitration agreement, the jury trial right vanishes.”¹⁴⁵

Interestingly, even with the historical importance of jury trials, arbitrators have been placed in a higher regard and are given more discretion than jurors. The Supreme Court has been clear on this issue by stating that the trial judge in the federal system “has . . . discretion to grant a new trial if the verdict appears . . . to be against the weight of the evidence.”¹⁴⁶ However, concerning arbitration, the Supreme Court has stated that because arbitrators do not have the benefit of judicial instruction on the law, “they need not give their reasons for their results; the record of their proceedings is not as complete as it is in a court trial; and judicial review of an award is more limited than judicial review of a trial.”¹⁴⁷

The language in the preceding sentence is startling—an unreasoned decision made by individuals with neither instruction on nor mandate to follow any applicable law is given less judicial review than a decision made by a group of individuals who are given instruction on the law and are mandated to follow and apply that law to the facts of a case.

Mandatory PDAs also force investors to give up their constitutional right to due process of law. In *Davis v. Prudential*, a customer arbitration case, the Eleventh Circuit held that “the state action element of a due process claim is absent in private arbitration cases.”¹⁴⁸ The Seventh Circuit similarly held in *Elmore v. Chicago & Illinois Midland Railway Co.* that “the fact that a private arbitrator denies the procedural safeguards that are encompassed by the term ‘due process of law’ cannot give rise to a constitutional complaint.”¹⁴⁹ It is important to note, that FINRA is a private entity. For the purposes of National Association of Securities Dealers (NASD) (now FINRA) arbitrations, courts have found that actions in the arbitration process are not state action because they arise from a private agreement.¹⁵⁰ FINRA may engage in quasi-judicial functions, but that does not mean it is a state actor.¹⁵¹

145. *Am. Heritage Life Ins. Co.*, 294 F.3d at 711 (quoting *Cremin v. Merrill Lynch, Pierce, Fenner & Smith, Inc.*, 957 F. Supp. 1460, 1471 (N.D. Ill. 1997)).

146. *Gasparini v. Ctr. for Humanities, Inc.*, 518 U.S. 415, 433 (1996) (quoting *Byrd v. Blue Ridge Rural Elec. Coop., Inc.*, 356 U.S. 525, 537 (1958)).

147. *Bernhardt v. Polygraphic Co. of Am.*, 350 U.S. 198, 203 (1956).

148. *Davis v. Prudential Securities, Inc.* 59 F.3d 1186, 1191–92 (11th Cir. 1995); see *Austern v. Chi. Bd. Options Exch., Inc.*, 716 F. Supp. 121, 125 (S.D.N.Y. 1989) (ruling that conduct of exchange arbitration panel was not state action), *aff’d*, 898 F.2d 882 (2d Cir. 1990).

149. *Elmore v. Chi. & Ill. Midland Ry. Co.*, 782 F.2d 94, 96 (7th Cir. 1986).

150. See *Bahr v. Nat’l Ass’n of Sec. Dealers*, 763 F. Supp. 584, 589 (S.D. Fla. 1991) (concluding that the conduct at issue was only of a private corporation, not a state actor).

151. See Board of Governors Decision, *supra* note 79, at 17 n.18 (citing *Desiderio v.*

It is instructive to look at an example of a mandatory PDAA in a brokerage contract under the construct above. For this Article, the mandatory PDAA at issue in the *FINRA v. Schwab* enforcement action will be used.¹⁵² First, the mandatory PDAA is not in the actual contract; it is provided as an additional disclosure document that does not need to be signed by the investor. The only mention of the mandatory PDAA is on Page 6 (out of 6) of the brokerage contract which states that “[b]y signing this Application, you acknowledge that you have received and read a copy of the attached Application Agreement, which contains a predispute arbitration provision” and “[t]he Agreement with Schwab includes a predispute arbitration clause. You acknowledge receipt of the predispute arbitration clause”¹⁵³ This language is interesting because it only speaks of a pre-dispute arbitration clause, but nothing about the *mandatory* nature of such clause.

Based upon the discussion above, investors can be forced to relinquish rights granted to them under the U.S. Constitution and forced to enter a dispute resolution system that is less transparent, with no guarantee of due process of law, and, as discussed below, it is infinitely more difficult to get judicial review of an erroneous decision.

2. Narrow Appellate Avenues Effectively Preclude Judicial Review of Arbitral Awards

For the past two decades, arbitration in forums sponsored by the securities industry has been the primary mechanism for the resolution of disputes among investors, brokerage firms, and brokers. FINRA currently operates as the largest dispute resolution forum in the securities industry and has created rules by which its arbitrators must operate.¹⁵⁴ It is

Nat'l Ass'n of Secs. Dealers, 191 F.3d 198, 206–07 (2d Cir. 1999) (affirming trial court's dismissal of plaintiff's constitutional claims challenging the arbitration clause in the Form U4 because NASD is not a state actor nor can its actions be fairly attributable to the state)); *see also Austern*, 716 F. Supp. at 125 (holding that the conduct of an arbitration panel “did not in any way constitute state action”); *FDIC v. Air Fla. Sys., Inc.*, 822 F.2d 833, 842 n.9 (9th Cir. 1987) (“The arbitration involved here was private, not state, action; it was conducted pursuant to contract by a private arbitrator. Although Congress, in the exercise of its commerce power, has provided for some governmental regulation of private arbitration agreements, we do not find in private arbitration proceedings the state action requisite for a constitutional due process claim.”).

152. *Accounts*, CHARLES SCHWAB, http://www.schwab.com/public/schwab/investing/accounts_products/accounts (last visited Mar. 26, 2015) (providing a link to the Brokerage Account Application Agreement).

153. *Id.* at 6.

154. *See* FINRA RULE 12000, FINRA MANUAL: CODE OF ARBITRATION PROCEDURE FOR CUSTOMER DISPUTES, FIN. INDUS. REGULATORY AUTH. (as amended in 2008), *available at*

important to note that this forum is separate from FINRA's disciplinary function discussed earlier in the Article. These rules cover filing and responding to a claim, hearings and discovery, and appointment and authority of the arbitrator(s).¹⁵⁵ FINRA's rules do not provide a process through which a party may challenge an arbitration award. Thus, to challenge an award a party must file an action in federal court to vacate the decision under the very limited circumstances prescribed in the FAA:

(1) where the award was procured by corruption, fraud, or undue means; (2) where there was evident partiality or corruption in the arbitrators, or either of them; (3) where the arbitrators were guilty of misconduct in refusing to postpone the hearing, upon sufficient cause shown, or in refusing to hear evidence pertinent and material to the controversy; or of any other misbehavior by which the rights of any party have been prejudiced; or (4) where the arbitrators exceeded their powers, or so imperfectly executed them that a mutual, final, and definite award upon the subject matter submitted was not made.¹⁵⁶

Parties who seek to challenge a final arbitration award have very few options. As noted above, there are limited avenues to vacate a valid arbitration award. To determine the scope of these options, it is useful to analyze the case law. Because the appellate rights originate under the FAA and not FINRA rules, the cases discussed below include non-securities arbitrations.

*a. Burden of Proof Upon Aggrieved Party and Courts
Extremely Deferential to Arbitrators*

As a starting point, a court's limited ability to review an arbitration award has been described as "among the narrowest known to the law."¹⁵⁷ There is a presumption that arbitration awards will be confirmed.¹⁵⁸

http://finra.complinet.com/en/display/display_main.html?rbid=2403&element_id=4096 (detailing the rules by which the FINRA arbitrators are bound).

155. *Id.*

156. 9 U.S.C. § 10 (2012 & Supp. I 2013).

157. *See, e.g., Denver & Rio Grande W. R.R. v. Union Pac. R.R.*, 119 F.3d 847, 849 (10th Cir. 1997) (quoting *ARW Exploration Corp. v. Aguirre*, 45 F.3d 1455, 1462 (10th Cir. 1995) (holding a court could not review an arbitration award despite having jurisdiction over the matter).

158. *See Frazier v. CitiFinancial Corp., LLC*, 604 F.3d 1313, 1324 (11th Cir. 2010) (holding that a court has no basis to overturn an arbitrator's decision outside of statutory grounds).

The party moving to vacate bears the burden of refuting every rational basis on which the arbitrator could have relied.¹⁵⁹ In *Collins v. Chicago Investment Group, LLC*, the plaintiff was appealing an adverse determination by a FINRA arbitration panel.¹⁶⁰ The U.S. District Court for the District of Nevada found that “[t]he burden of establishing grounds for vacating an arbitration award is on the party seeking vacatur”¹⁶¹ and that “[r]eview of the award is ‘both limited and highly deferential.’”¹⁶² In *Goldman Sachs Execution & Clearing, L.P. v. Official Unsecured Creditors’ Committee of Bayou Group, LLC*, the Second Circuit reinforced that the review of an arbitration award “is ‘highly deferential’ to the arbitrators, and relief on such a claim is therefore ‘rare.’”¹⁶³

The Second Circuit held in *Wallace v. Buttar* that the party moving to vacate an arbitration award bears the “heavy burden of showing that the award falls within a very narrow set of circumstances delineated by statute and case law.”¹⁶⁴ In *Banco de Seguros del Estado v. Mutual Marine Office, Inc.*, the same court stated it has:

‘consistently accorded the narrowest of readings’ to the FAA’s authorization to vacate awards . . . [and the Court’s] inquiry ‘focuses on whether the arbitrators had the power based on the parties’ submissions or the arbitration agreement, to reach a certain issue, not whether the arbitrators correctly decided that issue.’¹⁶⁵

159. *First Pres. Capital, Inc. v. Smith Barney, Harris Upham & Co., Inc.*, 939 F. Supp. 1559, 1564 (S.D. Fla. 1996).

160. *Collins v. Chi. Inv. Group, LLC*, 2012 U.S. Dist. LEXIS 37217 (D. Nev. 2012).

161. *Id.* at *1 (citing *U.S. Life Ins. Co. v. Superior Nat’l Ins. Co.*, 591 F.3d 1167, 1173 (9th Cir. 2010)).

162. *Id.* at *2 (quoting *Poweragent Inc. v. Elec. Data. Sys. Corp.*, 358 F.3d 1187, 1193 (9th Cir. 2004)).

163. *Goldman Sachs Execution & Clearing, L.P. v. Official Unsecured Creditors’ Comm. of Bayou Grp., LLC*, 491 Fed. App’x. 201, 203 (2d Cir. 2012) (quoting *STMicroelectronics, N.V. v. Credit Suisse Securities (USA) LLC*, 648 F.3d 68, 78 (2d Cir. 2011)).

164. *Wallace v. Buttar*, 378 F.3d 182, 189 (2d Cir. 2004) (quoting *Duferco Int’l Steel Trading v. T. Klaveness Shipping A/S*, 333 F.3d 383, 388 (2d Cir. 2003)).

165. *Banco de Seguros del Estado v. Mut. Marine Office, Inc.*, 344 F.3d 255, 262 (2d Cir. 2003) (quoting *Westerbeke Corp. v. Daihatsu Motor Co.*, 304 F.3d 200, 220 (2d Cir. 2002)); see *Stolt-Nielson S.A. v. AnimalFeeds Int’l Corp.*, 559 U.S. 662, 67 (2010) (stating that a party seeking vacatur of an arbitrator’s decision “must clear a high hurdle”).

The Court's review is to "determine only whether the arbitrator did his job—not whether he did it well, correctly, or reasonably, but simply whether he did it."¹⁶⁶

The courts' opinions concerning deference to arbitrators are in direct conflict with the opinions concerning deference to a governmental agency. For example, as noted above, courts are not obligated to defer to an agency's interpretation of Supreme Court precedent.¹⁶⁷ By contrast, arbitrators do not have to follow Supreme Court precedent, but, in the event they do, a court's review of an arbitration award is 'highly deferential' to the arbitrators. This contradiction reflects a decision to arbitrarily grant deference to a group of individuals who generally have less legal training and would ultimately be more likely to misinterpret Supreme Court jurisprudence.

b. Lack of Rationale Does Not Invalidate Award

FINRA rules do not require an arbitrator to provide a reasoned opinion for his or her decision.¹⁶⁸ In *Goldman Sachs*, the Court held that "an arbitral award must 'be enforced, despite a court's disagreement with it on the merits, if there is a *barely colorable justification* for the outcome reached.'"¹⁶⁹ Similarly, in *STMicroelectronics*, the Court held that even where an arbitration panel does "not explain the reason for [its] decision, we will uphold it if we can discern any valid ground for it."¹⁷⁰

The Ninth Circuit held in *Kyocera Corp. v. Prudential-Bache Trade Services, Inc.* that "[n]either erroneous legal conclusions nor unsubstantiated factual findings justify federal court review of an arbitral award under the [FAA], which is unambiguous in that regard."¹⁷¹ Then, in *D.H. Blair & Co., Inc. v. Gottdiener*, the Second Circuit held that the "arbitrator's rationale for an award need not be explained, and the award should be confirmed 'if a ground for the arbitrator's decision can be inferred from the facts of the case.' Only 'a barely colorable justification

166. *Mountaineer Gas Co. v. Oil, Chem. & Atomic Workers Int'l Union*, 76 F.3d 606, 608 (4th Cir. 1996).

167. *N.Y. N.Y., LLC v. N.L.R.B.*, 313 F.3d 585, 590 (D.C. Cir. 2002).

168. *Decisions and Awards*, FIN. INDUS. REGULATORY AUTH., available at <http://www.finra.org/arbitrationandmediation/arbitration/process/decisionawards/> (last visited Mar. 5, 2015).

169. *Goldman*, 491 Fed. App'x. at 204 (emphasis in original) (quoting *Wallace*, 378 F.3d at 190).

170. *STMicroelectronics, N.V. v. Credit Suisse Secs. (USA) LLC*, 648 F.3d 68, 78 (2d Cir. 2011).

171. *Kyocera Corp. v. Prudential-Bache Trade Servs., Inc.*, 341 F.3d 987, 994 (9th Cir. 2003).

for the outcome reached' by the arbitrators, is necessary to confirm the award."¹⁷²

Using an example from the securities industry, in *Raiford v. Merrill Lynch, Pierce, Fenner & Smith, Inc.*, the Eleventh Circuit held that an arbitration award for brokerage churning damages was not shown to be arbitrary and capricious where the arbitrator gave no explanation at all for his decision.¹⁷³ Even though in this case the decision favored the investor, the next case could just as easily favor the broker.

c. Impartiality, Bias of Arbitrators, and Presentation of Evidence

Under 9 U.S.C. § 10(a)(2), a party may vacate an arbitration award where there was evident partiality or corruption of the arbitrators. Courts have routinely held that this is a high burden to meet. For example, in *Scandinavian Reinsurance Co. Ltd. v. St. Paul Fire & Marine Insurance Co.*, the Second Circuit held that even where an arbitrator fails to abide by arbitral or ethical rules concerning disclosure of possible conflicts, such failure does not, in itself, entitle the losing party to vacatur of the arbitration award.¹⁷⁴ Indeed, the Sixth Circuit has noted that "[t]he alleged partiality must be direct, definite, and capable of demonstration."¹⁷⁵

Plaintiffs also have a very high burden for proving bias on the part of the arbitrators. For example, in *Kolel Beth Yechiel Mechil of Tartikov, Inc. v. YLL Irrevocable Trust*, an overheard conversation in which an arbitrator allegedly said that one party would receive a favorable decision was not direct or definite evidence of bias.¹⁷⁶

Courts have also held that arbitrators can proceed with a decision without necessarily having to procure all relevant evidence. In *United Paperworkers International Union, AFL-CIO v. Misco, Inc.*, the Supreme Court concluded that even assuming "that the arbitrator erred in refusing to consider the disputed evidence, his error was not in bad faith or so gross as

172. *D.H. Blair & Co., Inc. v. Gottdiener*, 462 F.3d 95, 110 (2d Cir. 2006) (quoting *Barbier v. Shearson Lehman Hutton, Inc.*, 948 F.2d 117, 121 (2d Cir. 1991); *Landy Michaels Realty Corp. v. Local 32B-32J, Serv. Emps. Int'l Union*, 954 F.2d 794, 797 (2d Cir. 1992)).

173. *Raiford v. Merrill Lynch, Pierce, Fenner & Smith, Inc.*, 903 F.2d 1410 (11th Cir. 1990).

174. *Scandinavian Reinsurance Co. Ltd. v. St. Paul Fire & Marine Ins. Co.*, 668 F.3d 60, 74 (2d Cir. 2012).

175. *Nationwide Mut. Ins. Co. v. Home Ins. Co.*, 278 F.3d 621, 626 (6th Cir. 2002) (quoting *Andersons, Inc. v. Horton Farms, Inc.*, 166 F.3d 308, 329 (6th Cir. 1998)).

176. *Kolel Beth Yechiel Mechil of Tartikov, Inc. v. YLL Irrevocable Trust*, 729 F.3d 99, 106 (2d Cir. 2013).

to amount to affirmative misconduct.”¹⁷⁷ Additionally, in *National Casualty Co. v. First State Insurance Group*, the First Circuit held that there was no misconduct when the arbitration panel issued the award without forcing the defendant company to produce relevant documents, because the arbitration panel acted within its authority when it chose to render a decision after drawing inferences against the company as to what documents would show.¹⁷⁸ In short, the arbitrators were allowed to divine what the documents would show without the necessity of reviewing them.

d. “*Manifest Disregard of the Law*”

Barring outright fraud, deceit, or impartiality of the arbitrators, claims to vacate an arbitration award must be brought under 9 U.S.C. § 10(a)(4), which states that an award may be vacated “where the arbitrators exceeded their powers, or so imperfectly executed them that a mutual, final, and definite award upon the subject matter submitted was not made.”¹⁷⁹ As noted above, the aggrieved party has a high burden to show that the arbitrator exceeded his or her power. In order to overturn an arbitral decision for excess use of power on the part of the arbitrators, the terms of the award must be “completely irrational.”¹⁸⁰ To be “completely irrational,” the arbitrators’ decision must “‘escape[] the bounds of rationality’” and be entirely unsupported by the record.¹⁸¹

Multiple circuits have also found two additional grounds upon which an award may be vacated: namely, where the arbitrator’s award is in “manifest disregard” of the law¹⁸² or the terms of the relevant agreement between the parties are outside the reasonable expectations of the weaker party.¹⁸³ Courts have found that “[m]anifest disregard of the law’ means

177. *United Paperworkers Int’l Union, AFL-CIO v. Misco, Inc.*, 484 U.S. 29, 40 (1987).

178. *Nat’l Cas. Co. v. First State Ins. Group*, 430 F.3d 492, 497 (1st Cir. 2005).

179. 9 U.S.C. § 10(a)(4) (2012 & Supp. I 2013).

180. *Sherrock Bros., Inc. v. DaimlerChrysler Motors Co.*, 260 F. App’x 497, 501 (3d Cir. 2008).

181. *Franko v. Ameriprise Fin. Servs.*, 2009 U.S. Dist. LEXIS 48907, at *10 (E.D. Pa. 2009) (quoting *Southco, Inc. v. Reell Precision Mfg. Corp.*, 556 F. Supp. 2d 505, 511 (E.D. Pa. 2008)).

182. This Article contends that the Supreme Court’s decision in *Hall Street Assocs., L.L.C. v. Mattel, Inc.* does not preclude “manifest disregard” of the law as a ground to vacate an arbitration award. *Comedy Club, Inc. v. Improv W. Assoc.*, 553 F.3d 1277, 1290 (9th Cir. 2009), *cert. denied*, *Improv W. Assoc. v. Comedy Club, Inc.*, 558 U.S. 824 (2009) (concluding that “after *Hall Street Associates*, manifest disregard of the law remains a valid ground for vacatur” because it is “shorthand for a statutory ground under the FAA, specifically 9 U.S.C. § 10(a)(4)”).

183. *See, e.g., Schwartz v. Merrill Lynch & Co., Inc.*, 665 F.3d 444, 451–52 (2d Cir. 2011) (holding that a court may set aside an arbitration award if it was rendered in manifest

something more than just an error in the law or a failure on the part of the arbitrators to understand or apply the law.”¹⁸⁴ In *Misco*, the Supreme Court held that “as long as the arbitrator is even arguably construing or applying the contract and acting within the scope of his authority, that a court is convinced he committed serious error does not suffice to overturn his decision.”¹⁸⁵ It is not enough to say that the arbitrator erred in resolving conflicting precedent in favor of the defendant.¹⁸⁶ “It must be clear from the record that the arbitrators recognized the applicable law and then ignored it.”¹⁸⁷ For a plaintiff to succeed, he or she is required to show that “the arbitrators were aware of the law, understood it correctly, found it applicable to the case before them, and yet chose to ignore it in propounding their decision.”¹⁸⁸

To provide an example of how narrowly courts view the “manifest disregard” of the law, in *Montes v. Shearson Lehman Brothers, Inc.*, the prevailing party of the arbitration contended that the arbitrator could ignore the law if doing so would provide a fair result.¹⁸⁹ The Eleventh Circuit vacated the award because:

the party who obtained the favorable award had conceded to the arbitration panel that its position was not supported by the law, which required a different result, and had urged the panel not to follow the law . . . [and] that blatant appeal to disregard the law was explicitly noted in the arbitration panel’s award.¹⁹⁰

It is important to note that the *Montes* case “remains the sole instance in which the 11th Circuit has ‘ever found the exceptional circumstances that

disregard of the law).

184. See, e.g., *Fin. Network Inv. Corp. v. Karoon*, 540 Fed. App’x 597 (9th Cir. 2013) (providing a working definition of manifest disregard of the law).

185. *United Paperworkers Int’l Union, AFL-CIO v. Misco, Inc.*, 484 U.S. 29, 38 (1987).

186. *Goldman v. Architectural Iron Co.*, 306 F.3d 1214, 1217 (2d Cir. 2002).

187. *Mich. Mut. Ins. Co. v. Unigard Sec. Ins. Co.*, 44 F.3d 826, 832 (9th Cir. 1995); see *O.R. Sec., Inc. v. Prof’l Planning Assocs.*, 857 F.2d 742, 747 (11th Cir. 1988) (stating that the knowing disregard of the law must be apparent on the face of the record.); see also *Carter v. Health Net of Cal., Inc.*, 374 F.3d 830, 838 (9th Cir. 2004) (quoting *Mich. Mut. Ins.*, 44 F.3d at 832) (noting that “[a]s federal courts of appeals have repeatedly held, “manifest disregard of the law” means something more than just an error in the law or a failure on the part of the arbitrators to understand or apply the law. It must be clear from the record that the arbitrators recognized the applicable law and then ignored it.”).

188. *Remmey v. PaineWebber, Inc.*, 32 F.3d 143, 149 (4th Cir. 1994).

189. *Montes v. Shearson Lehman Bros., Inc.*, 128 F.3d 1456, 1461–62 (11th Cir. 1997).

190. *Id.* at 1464 (Carnes, J., concurring).

satisfy the exacting requirements' justifying the vacation of an award due to manifest disregard for the law."¹⁹¹

Courts will not disturb factual findings by the arbitrator, if any are even made, and have consistently held that arbitration awards should not be vacated even if the arbitrators erred by making "contradictory findings of fact."¹⁹² For example, in *Oxford Health Plans, LLC v. Sutter*, the Supreme Court held that "convincing a court of an arbitrator's error—even his grave error—is not enough. So long as the arbitrator was 'arguably construing' the contract—which this one was—a court may not correct his mistakes under [the FAA]."¹⁹³ Thus, the court noted that "[w]hether or not the panel's findings are supported by the evidence in the record is beyond the scope of [the Court's] review."¹⁹⁴ The lack of rationale required under the FAA affects the Court's ability to review an arbitration award—"when arbitrators do not state their reasons for an award, 'it is nearly impossible for the court to determine whether they acted in disregard of the law.'"¹⁹⁵

Further confirming the oppressively narrow view taken by courts when reviewing arbitration awards, in *Reliastar Life Insurance Co. v. EMC National Life Co.*, the Second Circuit held that "as long as the arbitrator is even arguably construing or applying the contract and acting within the scope of his authority,' a court's conviction that the arbitrator has 'committed serious error' in resolving the disputed issue 'does not suffice to overturn his decision.'"¹⁹⁶ Indeed, courts will not vacate an award under

191. *Rueter v. Merrill Lynch, Pierce, Fenner & Smith, Inc.*, 440 F. Supp. 2d 1256, 1262 (D. Ala. 2006) (quoting *Montes*, 128 F.3d at 1461.).

192. *See, e.g., Bosack v. Soward*, 586 F.3d 1096, 1106 (9th Cir. 2009).

193. *Oxford Health Plans LLC v. Sutter*, 133 S.Ct. 2064, 2070 (2013).

194. *Lagstein v. Certain Underwriters at Lloyd's, London*, 607 F.3d 634, 642 (9th Cir. 2010) (quoting *Bosack*, 586 F.3d at 1106); *see Kennecott Utah Copper Corp. v. Becker*, 195 F.3d 1201, 1204 (10th Cir. 1999) (clarifying the manifest disregard standard by noting it applies only to conclusions of law, not to the arbitrator's factual findings, which are beyond review).

195. *Aetna Cas. & Sur. Co. v. Dravo Corp.*, No. 97-149, 1997 U.S. Dist. LEXIS 11648, at *4 (E.D. Pa. Aug. 1, 1997) (quoting *O.R. Sec., Inc. v. Prof'l Planning Assoc., Inc.*, 857 F.2d 742, 747 (11th Cir. 1988)).

196. *Reliastar Life Ins. Co. v. EMC Nat'l Life Co.*, 564 F.3d 81, 86 (2d Cir. 2009) (quoting *United Paperworkers Int'l Union AFL-CIO v. Misco, Inc.*, 484 U.S. 29, 38 (1987)). Where parent contended that stock in family corporation was placed in children's names for purposes of estate planning and tax avoidance, and that arbitrator improperly found that children owned majority of stock because they gave consideration for stock, arbitrator's failure to specifically discuss any legal reasoning does not establish that arbitrator identified applicable law and disregarded it, as required to vacate award based on manifest disregard of law. *Van Horn v. Van Horn*, 393 F. Supp. 2d 730 (N.D. Iowa 2005); *see Hutchinson v. Farm Family Cas. Ins. Co.*, 500 F. Supp. 2d 87 (D. Conn. 2007) (holding that arbitration award may not be vacated on the grounds that the arbitrator failed to interpret correctly law applicable to issues in dispute or misinterpreted underlying contract).

this standard “because of ‘a simple error in law or a failure by the arbitrators to understand or apply it’ but only when a party clearly demonstrates ‘that the panel intentionally defied the law.’”¹⁹⁷ Courts have also found that they will not “vacate an arbitral award merely because [they are] convinced that the arbitration panel made the wrong call on the law.”¹⁹⁸ Arbitrators “cannot be said to disregard a law that is unclear or not clearly applicable. Thus, misapplication of an ambiguous law does not constitute manifest disregard.”¹⁹⁹ Arbitrariness and capriciousness is also not an independent non-statutory ground for vacatur under the FAA.²⁰⁰

It is important to note that even though arbitrators have virtually unlimited discretion regarding the facts and the law in a specific case, they must still generally abide by FINRA’s procedural rules of arbitration.²⁰¹ In practice, however, because of the lack of reasoning required in the arbitration award and the courts’ hesitance to scrutinize arbitrations post-award, it is likely that there will be very few successful challenges using this argument.

e. Analysis

Based upon the case law discussed above, it appears that an investor or financial service provider seeking to vacate an arbitration award has virtually no chance of being successful in the absence of outright fraud or an instance where the arbitrator intentionally disregards the law and then publishes a written decision attesting to those facts.²⁰² It is telling that even adamant supporters of arbitration agree that there is virtually no effective way to appeal an erroneous arbitral award. Catholic University law professor Peter B. Rutledge, who has authored numerous studies on the benefits of arbitration,²⁰³ has stated that “the grounds for vacatur are

197. *StMicroelectronics, N.V. v. Credit Suisse Sec. (USA) LLC*, 648 F.3d 68, 78 (2d Cir. 2011) (quoting *Duferco Int’l Steel Trading v. T. Klaveness Shipping A/S*, 333 F.3d 383, 389, 393 (2d Cir. 2003)).

198. *Wallace v. Buttar*, 378 F.3d 182, 190 (2d Cir. 2004).

199. *Duferco*, 333 F.3d at 383.

200. *Brabham v. A.G. Edwards & Sons, Inc.*, 376 F.3d 377 (5th Cir. 2004).

201. *See Kashner Davidson Sec. Corp. v. Mscisz*, 531 F.3d 68 (1st Cir. 2008) (holding an arbitrator’s failure to abide by procedural rules when arriving at the arbitral award may support a manifest disregard of the law challenge.), *cf.* *Delta Mine Holding Co. v AFC Coal Proprs.*, 280 F.3d 815 (8th Cir. 2001) (determining that only the grounds listed under 9 U.S.C. § 10 justify vacating an award and arbitration rules and ethical codes do not have force of law).

202. *See supra* Section II.C.2.

203. *See, e.g.*, PETER B. RUTLEDGE, CHAMBER OF COMMERCE INST. FOR LEGAL REFORM, *ARBITRATION—A GOOD DEAL FOR CONSUMERS: A RESPONSE TO PUBLIC CITIZEN (2008)*, available at <http://www.adrforum.com/rcontrol/documents/ResearchStudiesAndStatistics/>

themselves extremely narrow, and the opportunity for judicial review of the award's substance virtually non-existent."²⁰⁴

To show how absurd these requirements are in practice, courts have found that "[n]either erroneous legal conclusions nor unsubstantiated factual findings justify federal court review"²⁰⁵ and there is no requirement for an arbitrator to provide a reasoned opinion for his or her decision. Based upon the legal precedent, an arbitrator may provide an incorrect legal conclusion, which is not substantiated by any factual finding, and the investor will still not be able to vacate the award.²⁰⁶

For example, assume that the investor wanted to vacate the arbitration award based upon the fact that the arbitrator found that his or her broker was not a fiduciary. Because the existence of a fiduciary duty is fact specific, the Court would not be able to review that determination. Even if the court did have the power to review the factual determination, the arbitrator has no legal duty to provide one. Even if the court believed that the arbitrator incorrectly interpreted the law by finding that the broker was not a fiduciary, unless the investor was able to show that the arbitrator knew the law and specifically ignored it, the court would not be able to vacate the arbitration.²⁰⁷

Investors may even be sanctioned for attempting to vacate the arbitration award. In *Rueter v. Merrill Lynch, Pierce, Fenner & Smith*, the District Court for the Northern District of Alabama denied the investor's motion to vacate and for reconsideration but granted the broker-dealer's motion for sanctions.²⁰⁸ The Court held that a sanctions award was proper because the investor's legal theory had no reasonable chance of success and could not be advanced as a reasonable argument to change existing law.²⁰⁹ Going further, even if an investor is able to vacate an adverse arbitration award, the investor "must then return to arbitration to relitigate the matter

200804ArbitrationGoodForConsumers-Rutledge.pdf (providing a detailed analysis of the benefits of arbitration).

204. Peter B. Rutledge, *Market Solutions to Market Problems: Re-Examining Arbitral Immunity as a Solution to Unfairness in Securities Arbitration*, 26 PACE L. REV. 113, 125 (2005).

205. *Kyocera Corp. v. Prudential-Bache Trade Servs., Inc.*, 341 F.3d 987, 994 (9th Cir. 2003).

206. *See, e.g., Oxford Health Plans LLC v. Sutter*, 133 S.Ct. 2064, 2070 (2013); *Bosack v. Soward*, 586 F.3d 1096, 1106 (9th Cir. 2009)

207. *Wachovia Sec., LLC v. Brand*, 671 F.3d 472, 483 (4th Cir. 2012) (determining that in order to vacate an arbitration award due to manifest disregard of the relevant law, the movant must show that: "(1) the applicable legal principle is clearly defined and not subject to reasonable debate; and (2) the arbitrator[] refused to heed that legal principle."").

208. *Rueter v. Merrill Lynch, Pierce, Fenner & Smith, Inc.*, 440 F. Supp. 2d 1256 (N.D. Ala. 2006).

209. *Id.*

at substantial expense, on top of the already substantial expense of bringing the vacatur action in court.”²¹⁰

This author is aware of the arguments that the finality of an arbitration award benefits the investor, focusing on the costs to an investor if the firm appeals the trial court’s decision, and the speed of arbitration panels versus the court system.²¹¹ Based upon the current legal structure governing arbitrations, however, an arbitration proceeding may ultimately end up costing the investor more than they would have paid at the court. This is because, as noted above, the investor must go to court to enforce a positive arbitration award or litigate an adverse arbitration award.²¹² Investors may also be sanctioned for appealing their awards.²¹³ The lack of appellate rights for either party in an arbitration is an absolute based upon rules and decisions that have been mandated by lower courts and confirmed by the Supreme Court.²¹⁴ So the real issue is whether the investor has a generally faster and cheaper dispute resolution or whether the investor has the ability to appeal an adverse arbitration award.

This argument is irrelevant, however, to the issue of mandatory PDAAS. It is important to note that while an arbitration *may* be less costly and more efficient, the aggrieved party will *definitely* not receive effective judicial review. In *Oxford Health Plans, LLC*, the Supreme Court stated that “the potential for . . . mistakes is the price of *agreeing* to arbitration.”²¹⁵ However, this is the exact choice that mandatory PDAAs take away from investors.

3. Unconscionability Analysis

This Article contends that mandatory PDAAs are *per se* unconscionable because they require investors to involuntarily waive certain constitutional rights and provide narrow appellate avenues that effectively preclude judicial review of arbitral awards. An agreement to arbitrate is valid, irrevocable, and enforceable, as a matter of federal law “save upon such grounds as exist at law or in equity for the revocation of

210. Peter B. Rutledge, *Toward a Contractual Approach for Arbitral Immunity*, 39 GA. L. REV. 151, 194 (2004).

211. See e.g., *The Securities Arbitration System: Hearing Before the Subcomm. on Capital Mkts., Ins. & Gov’t Sponsored Enters. of the H. Comm. On Fin. Servs.*, 109th Cong. 15-17 (2005) (statement of Marc E. Lackritz, President, Sec. Indus. Ass’n).

212. Rutledge, *supra* note 210, at 194.

213. See Reuter, 440 F. Supp. 2d 1256.

214. See, e.g., *Wilko v. Swan*, 346 U.S. 427, 436 (1953) (courts’ “power to vacate an award is limited”).

215. *Oxford Health Plans LLC v. Sutter*, 133 S. Ct. 2064, 2070 (2013) (emphasis added).

any contract.”²¹⁶ It is informative to look at how courts have interpreted the unconscionability doctrine and how it applies to the arguments made herein.

The Supreme Court in *Concepcion* held that “[a] finding of unconscionability requires ‘a “procedural” and a “substantive” element, the former focusing on “oppression” or “surprise” due to unequal bargaining power, the latter on “overly harsh” or “one-sided” results.’”²¹⁷ Most courts have focused on the unconscionability of the terms of the arbitration agreement rather than, as this Article contends, the legal regime governing arbitration. This Article contends that mandatory PDAs in the securities industry are *per se* unconscionable based upon the involuntary waiver of certain constitutional rights, lack of due process of law, and lack of judicial review.²¹⁸

Mandatory PDAs meet the first prong of the unconscionability test (procedural), because an investor would be surprised to learn that by agreeing to work with a financial advisor, he or she is required to relinquish his or her constitutional rights to access the courts and to due process of law.²¹⁹ As noted above in the Schwab brokerage agreement, the actual contract makes no mention of the mandatory nature of the arbitration; it is only when the investor reads the arbitration clause buried in the disclosure document that the investor realizes that he or she is subject to mandatory arbitration.²²⁰ Investors would also be surprised to learn that arbitrators do not have to follow the law in the arbitration proceeding, provide a reasoned decision, or gather all relevant evidence.²²¹ Investors are subject to oppression, because they generally do not realize they have involuntarily relinquished these rights until a dispute arises, which could be months or years after signing the initial contract.²²²

Mandatory PDAs also meet the second prong of the unconscionability test (substantive), because under the current legal regime, these agreements provide narrow appellate avenues that effectively preclude judicial review. As discussed above, the opportunity for judicial

216. *Moses H. Cone Mem’l Hosp. v. Mercury Constr. Corp.*, 460 U.S. 1, 24 (1983).

217. *AT&T Mobility LLC v. Concepcion*, 131 S. Ct. 1740, 1746 (2011) (quoting *Armendariz v. Found. Health Psychcare Servs.*, 6 P.3d 669, 690 (Cal. 2000)).

218. *See supra* Section II.C.1.

219. *Id.*

220. *See Schwab, supra* note 152 (exemplifying mandatory PDAA common in brokerage contracts).

221. *See supra* Section II.C.2.

222. *See* Jeff Sovern, et al., *'Whimsy Little Contracts' with Unexpected Consequences: An Empirical Analysis of Consumer Understanding of Arbitration Agreements* (St. John's Legal Studies Research Paper No. 14-0009, 2014), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2516432.

review of an arbitration award's substance is virtually non-existent.²²³ This Article contends that the legal regime governing the appeal of arbitration awards leads to extremely "one-sided" results; however, it is one-sided with regard to the prevailing party in the arbitration rather than the drafter of the agreement. For example, regardless of whether the investor or financial service provider prevail in the arbitration, the losing party effectively has no recourse to judicial review.²²⁴

In *Highway Equipment Co. v. Caterpillar, Inc.*, the Sixth Circuit held that "[t]o be unconscionable, a contract must be one 'which no man in his senses, not under delusion, would make, on the one hand, and which no fair and honest man could accept, on the other.'"²²⁵ Based upon the Section below that discusses how a mandatory PDAA would read under the current legal regime,²²⁶ no investor or financial service provider would be likely to sign such an agreement.

This Article also contends that this analysis does not run afoul of the holding in *Concepcion* that found that generally applicable contract defenses, such as duress or unconscionability, were alleged to have been applied in a fashion that disfavors arbitration.²²⁷ As noted above, arbitration is appropriate for a number of investors; however, investors should be able to make the choice based upon their own circumstances.

The unconscionability doctrine, as applied in this Article, only disfavors mandatory arbitration; it has no effect on voluntary arbitration. This is a distinction that has been repeatedly stated throughout this Article and has been explicitly affirmed by the courts—arbitration is a matter of *consent*.²²⁸ Once an investor is informed of his or her rights under an arbitration agreement and chooses to arbitrate his or her claim, the investor is subject to those terms. However, because of the constitutional rights involved and the lack of access to judicial review, the Supreme Court should clearly separate mandatory and voluntary arbitration. The court should use the substance of this Article to illustrate that while voluntary arbitration can be a less costly and efficient method for dispute resolution, forcing investors into that system will never be acceptable.

223. See *supra* Section II.C.2.

224. See *supra* Section II.C.2.e.

225. *Highway Equip. Co. v. Caterpillar, Inc.*, 908 F.2d 60, 65 (6th Cir. 1990) (quoting *Hume v. United States*, 131 U.S. 406, 410 (1889)).

226. See *infra* Section III.A.

227. *AT&T Mobility LLC v. Concepcion*, 131 S. Ct. 1740, 1746 (2011).

228. *Volt Info. Scis., Inc. v. Bd. of Trs.*, 489 U.S. 468, 479 (1989).

III. INFORMING THE INVESTOR

A. *How a Mandatory Pre-Dispute Arbitration Agreement Would Read Under the Current Legal Framework*

This section of the Article provides an example of what an investor's mandatory PDAA would look like if it clearly explained not only the arbitration process, but also the legal framework governing that process:

Mandatory Arbitration Provision

I agree that by working with you, I am being forced to submit all controversies that may arise between us to binding arbitration controlled by FINRA, the organization to which you are a member. Because such agreements are so pervasive throughout the securities industry, I have little choice but to accept this agreement if I wish to have assistance in reaching my financial goals. Such controversies include, but are not limited to, those involving any transaction in any of my accounts with you, or the construction, performance or breach of any agreement between us, whether entered into or occurring prior, on or subsequent to the date hereof.

I acknowledge that the arbitration process is generally more efficient and streamlined than litigation; however, by agreeing to work with you, I am waiving certain constitutional rights, including my right to seek remedies in court and my right to due process of law. I acknowledge that pre-arbitration discovery is more limited than the discovery allowed during a court proceeding and that I am not guaranteed a hearing to articulate my claims.²²⁹

If I disagree with the arbitrators' award, I acknowledge that there are few options for appealing the award.²³⁰ Even though awards must be in writing, arbitrators are not required to write reasoned opinions, provide explanations, or justify their decisions²³¹ and any ambiguity in the award

229. Jill I. Gross, *McMahon Turns Twenty: The Regulation of Fairness in Securities Arbitration*, 76 U. CIN. L. REV. 101, 108 (2008).

230. See *ARW Exploration Corp. v. Aguirre*, 45 F.3d 1455, 1462–63 (10th Cir. 1995) (explaining because of the courts' limited ability to review arbitration awards, their powers of review have been described as "among the narrowest known to the law."); see *Hall St. Assocs. v. Mattel, Inc.*, 552 U.S. 576 (2008) (establishing that contracting parties may not opt out of FAA's default standards for vacatur and fashion their own).

231. FINRA Code of Arbitration Procedure, Rule 12904 (2014); see *Atkinson v. Sinclair Refining Co.*, 370 U.S. 238 (1962) (holding the employer was not required to arbitrate its claim due to the contract); see also *Wilko v. Swan*, 346 U.S. 427, 436 (clarifying an arbitrators' "award may be made without explanation of their reasons and without a complete record of their proceedings . . .").

must be resolved, if possible, in a manner supporting confirmation of the award.²³²

Even if I can show that the arbitrators did not follow the arbitration (FINRA) rules, the Court will not overturn the award.²³³ Even if I can show that the arbitrators applied the incorrect law to my case, the Court will not overturn the award.²³⁴ Even if during the arbitration proceedings I overhear one of the arbitrators discussing his or her negative view of my case, the Court will not overturn the award.²³⁵ Even if the arbitrator fails to disclose possible conflicts of interest and is later revealed to have had previous business contacts with the opposing party, the Court will not overturn the award.²³⁶ Even if the arbitrators do not procure relevant documents that could have proven my case, the Court will not overturn the award.²³⁷ Even if I can show that the arbitrators' decision was not

232. See, e.g., *United Paperworkers Int'l Union v. Misco, Inc.*, 484 U.S. 29, 38 (1987) (determining “as long as the arbitrator is even arguably construing or applying the contract and acting within the scope of his authority, that a court is convinced he committed serious error does not suffice to overturn his decision.”); see *Atl. Aviation, Inc. v. EBM Grp., Inc.*, 11 F.3d 1276, 1282 (5th Cir. 1994) (indicating “courts should defer to the arbitrator’s resolution of the dispute whenever possible”).

233. See *Delta Mine Holding Co. v. AFC Coal Props., Inc.*, 280 F.3d 815, 820 (8th Cir. 2001) (explaining “only the statutory grounds in § 10(a) of the Act justify vacating an award; arbitration rules and ethical codes ‘do not have force of law.’”) (quoting *Merit Ins. Co. v. Leatherby Ins. Co.*, 714 F.2d 673, 680 (7th Cir. 1983)).

234. See, e.g., *Local 771, I.A.T.S.E. v. RKO Gen., Inc.*, 546 F.2d 1107, 1113 (2d Cir. 1977) (clarifying an arbitration award “will not be set aside for mistaken application of the law”); *Flexible Mfg. Sys. PTY v. Super Prods. Corp.* 86 F.3d 96, 100 (7th Cir. 1996) (highlighting that if “an arbitrator makes a mistake, by erroneously rejecting a valid, or even a dispositive legal defense, [it] does not provide grounds for vacating an award”); see also *Commercial Refrigeration, Inc. v. Layton Constr. Co.*, 319 F. Supp. 2d 1267 (D. Utah 2004) (explaining whether the arbitrator resolved legal issue correctly is not matter for court to determine); *Lagstein v. Certain Underwriters at Lloyd’s, London*, 607 F.3d 634, 641 (9th Cir. 2010) (clarifying that a “manifest disregard of the law” means something more than just an error in law or a failure on the part of an arbitrator to understand or apply law).

235. “Evidence of corruption must be abundantly clear in order to vacate” an arbitration award, and overheard conversation in which an arbitrator allegedly said that one party would receive a favorable decision was “not ‘direct’ or ‘definite’ evidence of bias.” *Kolel Beth Yechiel Mechil of Tartikov, Inc. v. YLL Irrevocable Trust*, 729 F.3d 99, 100 (2d Cir. 2013).

236. See *Gianelli Money Purchase Plan & Trust v. ADM Investor Servs., Inc.*, 146 F.3d 1309, 1309–10 (11th Cir. 1998) (explaining an arbitrator cannot be guilty of “evident partiality” by reason of “past business contacts between his employer” and interested party, “absent actual knowledge of a real or potential conflict”); *Sphere Drake Ins. Ltd. v. All Am. Life Ins. Co.*, 307 F.3d 617, 623 (7th Cir. 2002) (establishing an arbitrator’s “failure to make a full disclosure may sully his reputation for candor but does not demonstrate ‘evident partiality’ and thus does not spoil the award.”).

237. See, e.g., *Nat’l Cas. Co. v First State Ins. Group*, 430 F.3d 492, 498 (1st Cir. 2005) (explaining that where plaintiff sought to vacate arbitration award on grounds that the arbitration panel issued an award without forcing defendant company to produce relevant

supported by any evidence²³⁸ or that the decision was arbitrary,²³⁹ the Court will not overturn the award. I further acknowledge that even if I can show that the arbitrators made contradictory findings of fact, the Court will not overturn the award²⁴⁰ and that the only way to overturn an award on the ground of “manifest disregard of the law” is to show that the arbitrators recognized the applicable law and then chose to ignore it; however, since the arbitrators are not required to provide any explanation or reasons for their decision, I recognize that this is a near-impossible hurdle.²⁴¹

I acknowledge that I may be subject to sanctions for trying to vacate the arbitration award.²⁴² I further acknowledge that in the event I am successful in vacating an adverse arbitration award, I have to return to arbitration to re-litigate the matter at substantial expense, on top of the already substantial expense of bringing the vacatur action in court.²⁴³

In the event that I am granted an award by the arbitration panel, I acknowledge that even though FINRA rules state that “[a]ll monetary awards shall be paid within 30 days of receipt unless a motion to vacate has been filed with a court of competent jurisdiction,” FINRA has no power to enforce the award,²⁴⁴ other than to suspend or cancel the registration of the

documents, there was no misconduct because panel acted within its authority when it chose to render decision after drawing inferences against company as to what documents would show); *Scandinavian Reinsurance Co. v. St. Paul Fire & Marine Ins. Co.*, 668 F.3d 60, 77 (2d Cir. 2012) (holding “[t]he FAA does not bestow on a party the right to receive information about every matter that it might consider important or useful in presenting its case.”).

238. *See* *Coutee v. Barington Capital Grp., L.P.*, 336 F.3d 1128, 1134 (9th Cir. 2003) (affirming the courts “have no authority to re-weigh the evidence” presented to the arbitration panel).

239. *See* *Spungin v. GenSpring Family Offices, LLC*, 883 F. Supp. 2d 1193, 1198 (S.D. Fla. 2012) (citing *Hall St. Assocs., L.L.C. v. Mattel, Inc.*, 552 U.S. 576 (2008) (highlighting “it is apparent that an arbitration award cannot be vacated for being arbitrary and capricious.”); *see also* *Brabham v. A.G. Edwards & Sons, Inc.*, 376 F.3d 377, 382 (5th Cir. 2004) (establishing that “arbitrariness and capriciousness is not an accepted nonstatutory ground for vacatur” under the Federal Arbitration Act).

240. *See* *Bosack v. Soward*, 586 F.3d 1096, 1104 (9th Cir. 2009) (explaining “[w]hether or not the panel’s findings are supported by the evidence in the record is beyond the scope of our review.”).

241. *Dawahare v. Spencer*, 210 F.3d 666, 669 (6th Cir. 2000) (citing *Merrill Lynch, Pierce, Fenner & Smith, Inc. v. Jaros*, 70 F.3d 418, 421 (6th Cir.1995)) (clarifying “[a]rbitrators are not required to explain their decisions. If they choose not to do so, it is all but impossible to determine whether they acted with manifest disregard for the law.”).

242. *Rueter v. Merrill Lynch, Pierce, Fenner & Smith, Inc.*, 440 F. Supp. 2d 1256, 1265 (N.D. Ala. 2006).

243. *See supra* note 210 and accompanying text (emphasizing the return to arbitration if original award is overturned).

244. FINRA Code of Arbitration Procedure, Rule 12904 (2014); *see* *Fiero v. Fin. Indus. Regulatory Auth., Inc.*, 660 F.3d 569, 576 (2d Cir. 2011) (discussing the “seemingly inexplicable nature of a gap in the FINRA enforcement scheme: fines may be levied but not

2015]

TAKE IT OR LEAVE IT

615

broker or brokerage firm.²⁴⁵ I acknowledge that I must submit a petition to confirm the arbitration award with the courts.

This arbitration agreement provides the investor with a clear articulation of the arbitration process and the laws governing that process. This author contends that if an investor reads the language of the agreement, he or she would try to amend the language to exclude the arbitration provisions or just avoid working with the advisor altogether. With the pervasive use of mandatory PDAAs in the securities industry, however, investors may find this to be an insurmountable obstacle.

IV. POSSIBLE SOLUTIONS

This section of the Article offers possible solutions for protecting investors from being forced into a dispute resolution system that they never agreed to. The solutions include ending mandatory arbitration in brokerage and advisory contracts, or providing investors with a separate disclosure document to allow the advisor to receive an investor's informed consent and allowing the investor to opt-out of arbitration. This author was heartened by the recent news that FINRA has established an arbitration task force to "consider possible enhancements to its arbitration forum to improve the transparency, impartiality and efficiency of FINRA's securities arbitration forum for all participants."²⁴⁶

A. *Ending Mandatory Arbitration*

The most appropriate solution to this problem is to end the practice of placing mandatory PDAAs in brokerage and advisory contracts. As stated throughout this Article, arbitration in the securities industry is not necessarily bad, and many investors would benefit from arbitration; however, investors should be given this choice.

The issue of mandatory PDAAs, in a variety of industries, has been a contested issue in Congress since the mid 1990's. Since 1995, there have been 139 anti-arbitration bills introduced in Congress; however, none of these were passed into law.²⁴⁷ These include bills regarding brokerage

collected.").

245. FINRA Code of Arbitration Procedure, Rule 9554 (2014).

246. Press Release, Fin. Indus. Regulatory Auth., FINRA Announces Arbitration Task Force (July 17, 2014), available at <http://www.finra.org/Newsroom/NewsReleases/2014/P554192>.

247. Thomas V. Burch, *Regulating Mandatory Arbitration*, 2011 UTAH L. REV. 1309, 1310.

contracts, investor contracts, lending contracts, employment contracts, and many more.²⁴⁸

Some members of Congress have decided that they need to take action against these clauses in the securities industry, which they argue are unfair to investors who must forfeit their right to “their day in court” in order to obtain the professional services of a broker or investment adviser. In an interview with *InvestmentNews*, Rep. Keith Ellison (D-MN), a member of the House Financial Services Committee, stated that “[i]nvestors shouldn’t have to sign away their rights in order to work with a financial adviser or broker dealer to build a secure retirement. By removing some of the unfair advantages, consumers will be more eager to invest, which will create jobs and strengthen the economy.”²⁴⁹

Rep. Ellison introduced the Investor Choice Act of 2013 to address pre-dispute mandatory arbitration in brokerage and advisory agreements.²⁵⁰ The purpose of the bill is to “amend the Securities Exchange Act of 1934 to prohibit mandatory pre-dispute arbitration agreements, and for other purposes.”²⁵¹ In summary, the bill would make it unlawful for brokers, advisers, and certain other financial services professionals to mandate pre-dispute arbitration of claims and to prevent investors from seeking redress through class action lawsuits.²⁵²

It is important to note that even though Congress is looking at the issue of mandatory PDAAS in the securities industry, the SEC already has the authority to mandate that individuals and companies no longer use these clauses in their contracts. Section 921 of the Dodd-Frank Wall Street Reform and Consumer Protection Act (“Dodd-Frank Act”) authorized the SEC to adopt regulations to ban, condition, or limit mandatory arbitration clauses.²⁵³ This also allows FINRA to amend its rules concerning mandatory arbitration, because the SEC must approve any FINRA proposed rule. As of the date of this Article, no action has been taken.

This inaction on the part of the SEC is surprising, especially since over thirty years ago, in 1979, the SEC “expressed concern about the use of arbitration clauses in broker-dealer customer agreements which purport to

248. *See id.* (including appendix listing all 139 bills and a summary of the bills and their status).

249. Mark Schoeff, *House Democrat Introduces Bill to End Mandatory Arbitration*, INVESTMENTNEWS, (Aug. 5, 2013, 2:59 PM), <https://www.investmentnews.com/article/20130805/FREE/130809965?template=printart>.

250. Investor Choice Act of 2013, H.R. 2998, 113th Cong. (2013).

251. *Id.* at 1.

252. *Id.* at 1–2.

253. Dodd-Frank Wall Street Reform and Consumer Protection Act, Pub. L. No. 111-203, § 921, 124 Stat. 1376, 1841 (2010).

bind customers to arbitrate all future disputes with a broker-dealer.”²⁵⁴ In the Release, the SEC explicitly stated that:

[t]he Commission believes that the use of clauses which purport to bind customers to arbitrate all future disputes which may arise between them and their broker-dealers, without specifying the meaning, effect and enforceability thereof, *is inconsistent with just and equitable principles of trade*, and may raise serious questions of compliance with the anti-fraud provisions of the federal securities laws.²⁵⁵

The Senate has also recently asked the SEC to investigate the use of mandatory PDAA's in their most recent appropriations bill: “[t]he Committee directs that the SEC prioritize, initiate, and complete critical rulemakings to address . . . limiting use of mandatory arbitration agreements in brokerage customer service agreements . . .”²⁵⁶ This request reflects Congressional interest in understanding the impact of PDAA's in the financial services industry.

B. Informed Consent and Opt-Out Provision

The second option is to mandate that financial service providers obtain investors' informed consent before signing the advisory or brokerage agreement and allowing the investor to “opt-out” of the pre-dispute arbitration agreement. This can be handled by two different methods. The first method is for the financial service provider to furnish the investor with an arbitration document, separate from the contract, which describes the arbitration process (including options for appeal) at the beginning of the client relationship. The second method is for the financial service provider to deliver this document to the investor once a dispute arises and, at that time, the investor can make an informed decision whether to enter into arbitration. Both of these methods require an investor to choose a single method of dispute resolution. For example, an investor cannot choose litigation, but then ask for arbitration if the investor is losing in court.

Under this method, investors have the choice of arbitration or litigation. Currently, investors are generally notified by a clause buried in the boilerplate language of a brokerage or advisory contract or lengthy

254. Notice to Broker-Dealers Concerning Clauses in Customer Agreements which Provide for Arbitration of Future Disputes, Exchange Act Release No. 15,984, 1979 WL 174165, at *1 (July 2, 1979).

255. *Id.* (emphasis added).

256. 2015 APPROPRIATIONS BILL, *supra* note 5, at 132.

disclosure document. By requiring informed consent, an investor will be provided with all the information he or she needs to make an informed decision and will not have to suffer due to a clause, which he or she was not aware of, buried in a contract or long disclosure document.

CONCLUSION

An overwhelming majority of retail brokerage, and many investment advisory agreements, include language requiring that all disputes between the customer and the broker-dealer/investment adviser be resolved through arbitration, and only in rare instances can an investor open either a brokerage or investment advisory account without agreeing to submit to mandatory pre-dispute arbitration.

This Article has provided ample evidence that mandatory PDAAAs force investors to relinquish certain constitutional rights. This Article has also provided ample evidence that there is effectively no appellate review of arbitration awards and that it is virtually impossible for an investor to overturn an erroneous award. As noted above, some investor claims are appropriate for arbitration; however, the choice should belong to the investor. The SEC has the authority to ban these types of agreements; however, no action has been taken. Until these types of agreements are banned from the industry, investors will continue to be forced into a dispute resolution system that lacks transparency, requires the investor to relinquish certain constitutional rights, and lacks any effective mechanism to correct erroneous decisions.

77

Are Investors Reluctant to Realize Their Losses?

TERRANCE ODEAN*

ABSTRACT

I test the disposition effect, the tendency of investors to hold losing investments too long and sell winning investments too soon, by analyzing trading records for 10,000 accounts at a large discount brokerage house. These investors demonstrate a strong preference for realizing winners rather than losers. Their behavior does not appear to be motivated by a desire to rebalance portfolios, or to avoid the higher trading costs of low priced stocks. Nor is it justified by subsequent portfolio performance. For taxable investments, it is suboptimal and leads to lower after-tax returns. Tax-motivated selling is most evident in December.

THE TENDENCY TO HOLD LOSERS too long and sell winners too soon has been labeled the disposition effect by Shefrin and Statman (1985). For taxable investments the disposition effect predicts that people will behave quite differently than they would if they paid attention to tax consequences. To test the disposition effect, I obtained the trading records from 1987 through 1993 for 10,000 accounts at a large discount brokerage house. An analysis of these records shows that, overall, investors realize their gains more readily than their losses. The analysis also indicates that many investors engage in tax-motivated selling, especially in December. Alternative explanations have been proposed for why investors might realize their profitable investments while retaining their losing investments. Investors may rationally, or irrationally, believe that their current losers will in the future outperform their current

*University of California, Davis. This paper is based on my dissertation at the University of California, Berkeley. I would like to thank an anonymous referee, Brad Barber, Peter Klein, Hayne Leland, Richard Lyons, David Modest, John Nofsinger, James Poterba, Mark Rubinstein, Paul Ruud, Richard Sansing, Richard Thaler, Brett Trueman, and participants at the Berkeley Program in Finance, the NBER behavioral finance meeting, the Financial Management Association Conference, the American Finance Association meetings, and seminar participants at UC Berkeley, the Yale School of Management, the University of California, Davis, the University of Southern California, the University of North Carolina, Duke University, the Wharton School, Stanford University, the University of Oregon, Harvard University, the Massachusetts Institute of Technology, the Amos Tuck School, the University of Chicago, the University of British Columbia, Northwestern University, the University of Texas, UCLA, the University of Michigan, and Columbia University for helpful comments. I would also like to thank Jeremy Evnine and especially the discount brokerage house that provided the data necessary for this study. Financial support from the Nasdaq Foundation is gratefully acknowledged.

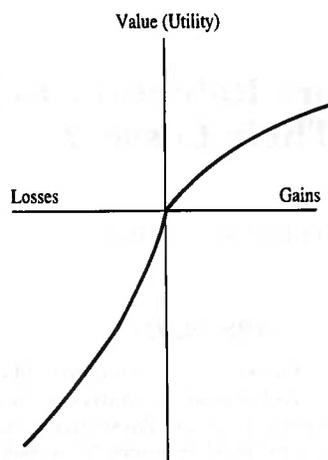


Figure 1. Prospect theory value function.

winners. They may sell winners to rebalance their portfolios. Or they may refrain from selling losers due to the higher transactions costs of trading at lower prices. I find, however, that when the data are controlled for rebalancing and for share price, the disposition effect is still observed. And the winning investments that investors choose to sell continue in subsequent months to outperform the losers they keep.

The next section of the paper discusses the disposition effect and literature related to it. Section II describes the data set and Section III describes the empirical study and its findings. Section IV discusses these findings and Section V concludes.

I. The Disposition Effect

A. Prospect Theory

The disposition effect is one implication of extending Kahneman and Tversky's (1979) prospect theory to investments. Under prospect theory, when faced with choices involving simple two and three outcome lotteries, people behave as if maximizing an "S"-shaped value function (see Figure 1). This value function is similar to a standard utility function except that it is defined on gains and losses rather than on levels of wealth. The function is concave in the domain of gains and convex in the domain of losses. It is also steeper for losses than for gains, which implies that people are generally risk-averse. Critical to this value function is the reference point from which gains and losses are measured. Usually the status quo is taken as the reference point; however, "there are situations in which gains and losses are

coded relative to an expectation or aspiration level that differs from the status quo.... A person who has not made peace with his losses is likely to accept gambles that would be unacceptable to him otherwise" (Kahneman and Tversky (1979)).

For example, suppose an investor purchases a stock that she believes to have an expected return high enough to justify its risk. If the stock appreciates and the investor continues to use the purchase price as a reference point, the stock price will then be in a more concave, more risk-averse, part of the investor's value function. It may be that the stock's expected return continues to justify its risk. However, if the investor somewhat lowers her expectation of the stock's return, she will be likely to sell the stock. What if, instead of appreciating, the stock declines? Then its price is in the convex, risk-seeking, part of the value function. Here the investor will continue to hold the stock even if its expected return falls lower than would have been necessary for her to justify its original purchase. Thus the investor's belief about expected return must fall further to motivate the sale of a stock that has already declined than one that has appreciated. Similarly, consider an investor who holds two stocks. One is up; the other is down. If the investor is faced with a liquidity demand, and has no new information about either stock, she is more likely to sell the stock that is up.

Throughout this study, investors' reference points are assumed to be their purchase prices. Though the results presented here appear to vindicate that choice, it is likely that for some investments, particularly those held for a long time over a wide range of prices, the purchase price may be only one determinant of the reference point. The price path may also affect the level of the reference point. For example, a homeowner who bought her home for \$100,000 just before a real-estate boom and had the home appraised for \$200,000 after the boom may no longer feel she is "breaking even" if she sells her home for \$100,000 plus commissions. If purchase price is a major component, though not the sole component, of reference point, it may serve as a noisy proxy for the true reference point. Using the proxy in place of the true reference point will make a case for the disposition effect more difficult to prove. It seems likely that if the true reference point were available the statistical evidence reported here would be even stronger.

B. An Alternative Behavioral Theory

Investors might choose to hold their losers and sell their winners not because they are reluctant to realize losses but because they believe that today's losers will soon outperform today's winners. If future expected returns for the losers are greater than those for the winners, the investors' belief would be justified and rational. If, however, future expected returns for losers are not greater than those for winners, but investors continue to believe

they are despite persistent evidence to the contrary, this belief would be irrational. In experimental settings Andreassen (1988) finds that subjects buy and sell stocks as if they expect short-term mean reversion.¹

Most of the analysis presented here does not distinguish between prospect theory and an irrational belief in mean reversion as possible explanations for why investors hold losers and sell winners. It may be that investors themselves do not always make a clear distinction. For example, an investor who will not sell a stock for a loss might convince himself that the stock is likely to bounce back rather than admit his unwillingness to accept a loss.

C. Taxes

Investors' reluctance to realize losses is at odds with optimal tax-loss selling for taxable investments. For tax purposes investors should postpone taxable gains by continuing to hold their profitable investments. They should capture tax losses by selling their losing investments, though not necessarily at a constant rate. Constantinides (1984) shows that when there are transactions costs, and no distinction is made between the short-term and long-term tax rates (as is approximately the case from 1987 to 1993 for U.S. federal taxes²), investors should gradually increase their tax-loss selling from January to December. Dyl (1977), Lakonishok and Smidt (1986), and Badrinath and Lewellen (1991) report evidence that investors do sell more losing investments near the end of the year.

Shefrin and Statman (1985) propose that investors choose to sell their losers in December as a self-control measure. They reason that investors are reluctant to sell for a loss but recognize the tax benefits of doing so. The end of the year is the deadline for realizing these losses. So each year, investors postpone realizing losses until December when they require themselves to sell losers before the deadline passes.

A sophisticated investor could reconcile tax-loss selling with her aversion to realize losses through a tax-swap. By selling her losing stock and purchasing a stock with similar risk characteristics, she could realize a tax-loss while maintaining the same risk exposure. Thaler (1985) argues that

¹ Subjects' tendencies to trade as if making regressive predictions diminish when their attention is focused on price changes rather than price levels (Andreassen (1988)) and when casual attributions for price trends, such as might normally be provided by the media, are made available (Andreassen (1987, 1990)).

² Prior to 1987 long-term capital gains tax rates were 40 percent of the short-term capital gains tax rates; from 1987 to 1993 long-term and short-term gains were taxed at the same marginal rates for lower income taxpayers. The maximum short-term rate at times exceeded the maximum long-term rate. In 1987 the maximum short-term rate was 38.5 percent and the maximum long-term rate was 28 percent. From 1988 to 1990 the highest income taxpayers paid a marginal rate of 28 percent on both long-term and short-term gains. In 1991 and 1992 the maximum long-term and short-term rates were 28 percent and 31 percent. In 1993 the maximum long-term and short-term rates were 28 percent and 39.6 percent.

people tend to segregate different gambles into separate mental accounts. These are then evaluated separately for gains and losses. A tax-swap requires closing such an account for a loss, which people are reluctant to do.

D. Previous Studies

Previous research³ offers some support for the hypothesis that investors sell winners more readily than losers, but this research is generally unable to distinguish among various motivations investors might have for doing so. Investors may be behaviorally motivated to hold losers and sell winners, that is, they may have value functions like those described in prospect theory or they may incorrectly expect mean-reverting prices. There are also rational reasons why investors may choose to hold their losers and sell their winners: (1) Investors who do not hold the market portfolio may respond to large price increases by selling some of the appreciated stock to restore diversification to their portfolios (Lakonishok and Smidt (1986)); (2) Investors who purchase stocks on favorable information may sell if the price goes up, rationally believing that price now reflects this information, and may continue to hold if the price goes down, rationally believing that their information is not yet incorporated into price (Lakonishok and Smidt (1986)); and (3) Because trading costs tend to be higher for lower priced stocks, and because losing investments are more likely to be lower priced than winning investments, investors may refrain from selling losers simply to avoid the higher trading costs of low-priced stocks (Harris (1988)).

The contribution of this paper is to demonstrate, with market data, that a particular class of investors (those with discount brokerage accounts) sell winners more readily than losers. Even when the alternative rational motivations listed above are controlled for, these investors continue to prefer selling winners and holding losers. Their behavior is consistent with prospect theory; it is also consistent with a (mistaken) belief that their winners and losers will mean revert.

³ Starr-McCluer (1995) finds that 15 percent of the stock-owning households interviewed in the 1989 and 1992 Surveys of Consumer Finances have paper losses of 20 percent or more. She estimates that in the majority of cases the tax advantages of realizing these losses would more than offset the trading costs and time costs of doing so. Heisler (1994) documents loss aversion in a small sample of futures speculators. In a study of individual federal tax returns, Poterba (1987) finds that although many investors do offset their capital gains with losses, more than 60 percent of the investors with gains or losses realized only gains. Weber and Camerer (1995) report experimental evidence of the disposition effect. Lakonishok and Smidt (1986) and Ferris, Haugen, and Makhija (1988) find a positive correlation between price change and volume. Bremer and Kato (1996) find the same correlation for Japanese stocks. Such a correlation could be caused by investors who prefer to sell winners and hold losers, but it could also be the result of buyers' trading preferences.

II. The Data

The data for this study are provided by a nationwide discount brokerage house. From all accounts active in 1987 (those with at least one transaction), 10,000 customer accounts are randomly selected. The data are in three files: a trades file, a security number to CUSIP file, and a positions file. Only the first two files are used in this study. The trades file includes the records of all trades made in the 10,000 accounts from January 1987 through December 1993. This file has 162,948 records, each record is made up of an account identifier, the trade date, the brokerage house's internal number for the security traded, a buy-sell indicator, the quantity traded, the commission paid, and the principal amount. Multiple buys or sells of the same stock, in the same account, on the same day, are aggregated. The security number to CUSIP table translates the brokerage house's internal numbers into CUSIP numbers. The positions file contains monthly position information for the 10,000 accounts from January 1988 through December 1993. Each of its 1,258,135 records is made up of the account identifier, year, month, internal security number, equity, and quantity. Accounts that were closed between January 1987 and December 1993 are not replaced; thus the data set may have some survivorship bias in favor of more successful investors. The data do not distinguish different account types. Therefore it is not possible to separate taxable accounts from tax-free accounts. Given the large sample size, we can expect the sample proportions of different account types to be close to the proportions for all of the brokerage's accounts. At the beginning of the data period, 20 percent of the brokerage's accounts were either IRA or Keogh accounts, and these accounts were responsible for 17.5 percent of all trades. The inclusion of these tax-exempt accounts will reduce tax-motivated trading in the data set, but with 80 percent of the accounts taxable, tax-motivated selling is easily detectable.

There are two data sets similar to this one described in the literature. Schlarbaum et al. (1978) and others analyze trading records for 2500 accounts at a large retail brokerage house for the period January 1964 to December 1970; Badrinath and Lewellen (1991) and others analyze a second data set provided by the same retail broker for 3000 accounts over the period January 1971 to September 1979. The data set studied here differs from these primarily in that it is more recent and comes from a discount broker. By examining discount brokerage records I can rule out the retail broker as an influence on observed trading patterns.

Badrinath and Lewellen (1991) look for evidence of tax-motivated trading and find that the ratio of stocks sold for a loss to those sold for a gain rises as the year progresses. Using a somewhat different measure, I also find evidence that investors increase their tax-motivated selling as the year progresses. However the focus of this paper, unlike that of Badrinath and Lewellen, is to test the disposition effect. As the next section describes, this is done by analyzing the rates at which investors realize gains and losses relative to their opportunities to do so.

III. Empirical Study

A. Methodology

This study tests whether investors sell their winners too soon and hold losers too long. It also investigates tax-motivated trading in December. To determine whether investors sell winners more readily than losers, it is not sufficient to look at the number of securities sold for gains versus the number sold for losses. Suppose investors are indifferent to selling winners or losers. Then in an upward-moving market they will have more winners in their portfolios and will tend to sell more winners than losers even though they had no preference for doing so.⁴ To test whether investors are disposed to selling winners and holding losers, we must look at the frequency with which they sell winners and losers relative to their opportunities to sell each.

By going through each account's trading records in chronological order, I construct for each date a portfolio of securities for which the purchase date and price are known. Clearly this portfolio represents only part of each investor's total portfolio. In most accounts there will be securities that were purchased before January 1987 for which the purchase price is not available, and investors may also have other accounts that are not part of the data set. Though the portfolios constructed from the data set are only part of each investor's total portfolio, it is unlikely that the selection process will bias these partial portfolios toward stocks for which investors have unusual preferences for realizing gains or losses.

I obtain information on splits and dividends as well as other price data needed for this study from the 1993 Center for Research in Security Prices daily stock file for NYSE, AMEX, and Nasdaq stocks. The study is limited to stocks for which this information is available. Of the 10,000 accounts, 6,380 trade stocks in the CRSP file for a total of 97,483 transactions.

Each day that a sale takes place in a portfolio of two or more stocks, I compare the selling price for each stock sold to its average purchase price to determine whether that stock is sold for a gain or a loss. Each stock that is in that portfolio at the beginning of that day, but is not sold, is considered to be a paper (unrealized) gain or loss (or neither). Whether it is a paper gain or loss is determined by comparing its high and low price for that day (as obtained from CRSP) to its average purchase price. If both its daily high and low are above its average purchase price it is counted as a paper gain; if they are both below its average purchase price it is counted as a paper loss; if its average purchase price lies between the high and the low, neither a gain or loss is counted. On days when no sales take place in an account, no gains or losses, realized or paper, are counted.

⁴ In Badrinath and Lewellen (1991) 49 percent of all round-trip sales are for a loss. In my database only 43 percent of such sales are for a loss. The difference could be due to different trading practices by retail and discount investors, but quite likely it simply reflects the greater rise in prices during the period I examine.

Suppose, for example, that an investor has five stocks in his portfolio, A, B, C, D, and E. A and B are worth more than he paid for them; C, D, and E are worth less. Another investor has three stocks F, G, and H in her portfolio. F and G are worth more than she paid for them; H is worth less. On a particular day the first investor sells shares of A and of C. The next day the other investor sells shares of F. The sales of A and F are counted as realized gains. The sale of C is a realized loss. Since B and G could have been sold for a profit but weren't, they are counted as paper gains. D, E, and H are paper losses. So for these two investors over these two days, two realized gains, one realized loss, two paper gains, and three paper losses are counted. Realized gains, paper gains, realized losses, and paper losses are summed for each account and across accounts. Then two ratios are calculated:

$$\frac{\text{Realized Gains}}{\text{Realized Gains} + \text{Paper Gains}} = \text{Proportion of Gains Realized (PGR)} \quad (1)$$

$$\frac{\text{Realized Losses}}{\text{Realized Losses} + \text{Paper Losses}} = \text{Proportion of Losses Realized (PLR)} \quad (2)$$

In the example $\text{PGR} = 1/2$ and $\text{PLR} = 1/4$. A large difference in the proportion of gains realized (PGR) and the proportion of losses realized (PLR) indicates that investors are more willing to realize either gains or losses.

Any test of the disposition effect is a joint test of the hypothesis that people sell gains more readily than losses and of the specification of the reference point from which gains and losses are determined. Some possible choices of a reference point for stocks are the average purchase price, the highest purchase price, the first purchase price, or the most recent purchase price. The findings of this study are essentially the same for each choice; results are reported for average purchase price. Commissions and dividends may or may not be considered when determining reference points, and profits and losses. Although investors may not consider commissions when they remember what they paid for a stock, commissions do affect capital gains and losses. And because the normative standard to which the disposition effect is being contrasted is optimal tax-motivated selling, commissions are added to the purchase price and deducted from the sales price in this study except where otherwise noted. Dividends are not included when determining which sales are profitable because they do not affect capital gains and losses for tax purposes. The primary finding of the paper, that investors are reluctant to sell their losers and prefer to sell winners, is unaffected by the inclusion or exclusion of commissions or dividends. In determining whether the stocks that are not sold on a particular day could have been sold for a gain or a loss, the commission for the potential sale is assumed to be the average commission per share paid when the stock was purchased.⁵ All gains and losses are calculated after adjusting for splits.

⁵ If, for potential sales, the commission is instead assumed to be the same percentage of principal as paid when the stock was purchased, the results do not significantly change.

Table I
PGR and PLR for the Entire Data Set

This table compares the aggregate Proportion of Gains Realized (PGR) to the aggregate Proportion of Losses Realized (PLR), where PGR is the number of realized gains divided by the number of realized gains plus the number of paper (unrealized) gains, and PLR is the number of realized losses divided by the number of realized losses plus the number of paper (unrealized) losses. Realized gains, paper gains, losses, and paper losses are aggregated over time (1987–1993) and across all accounts in the data set. PGR and PLR are reported for the entire year, for December only, and for January through November. For the entire year there are 13,883 realized gains, 79,658 paper gains, 11,930 realized losses, and 110,348 paper losses. For December there are 866 realized gains, 7,131 paper gains, 1,555 realized losses, and 10,604 paper losses. The *t*-statistics test the null hypotheses that the differences in proportions are equal to zero assuming that all realized gains, paper gains, realized losses, and paper losses result from independent decisions.

	Entire Year	December	Jan.–Nov.
PLR	0.098	0.128	0.094
PGR	0.148	0.108	0.152
Difference in proportions	-0.050	0.020	-0.058
<i>t</i> -statistic	-35	4.3	-38

There are two hypotheses to be tested. The first is that investors tend to sell their winners and hold their losers. Stated in terms of realization rates for gains and losses this is:

HYPOTHESIS 1: *Proportion of Gains Realized > Proportion of Losses Realized (for the entire year).*

The null hypothesis in this case is that $PGR \leq PLR$. The second hypothesis is that in December investors are more willing to sell losers and less willing to sell winners than during the rest of the year. That is:

HYPOTHESIS 2: *Proportion of Losses Realized – Proportion of Gains Realized in December > Proportion of Losses Realized – Proportion of Gains Realized in January–November.*

The null hypothesis here is: $PLR - PGR$ in December \leq $PLR - PGR$ in January through November.

B. Results

Table I reports the PGR realized and the PLR realized for the entire year, for January through November, and for December. We see that for the entire year investors do sell a higher proportion of their winners than of their losers. For both Hypothesis 1 and Hypothesis 2 the null hypotheses can be rejected with a high degree of statistical significance. A one-tailed test of the first null hypothesis, $PGR \leq PLR$, is rejected with a *t*-statistic greater than 35. The second null hypothesis, $PLR - PGR$ in December \leq $PLR - PGR$ in

January through November, is also rejected (t equals 16). These tests count each sale for a gain, sale for a loss, paper gain on the day of a sale, and paper loss on the day of a sale as separate independent observations.⁶ These observations are aggregated across investors. This independence assumption will not hold perfectly. For example, suppose an investor chooses not to sell the same stock on repeated occasions. It is likely that the decision not to sell on one date is not independent of the decision not to sell on another date. Alternatively, two investors may be motivated to sell the same stock on, or about, the same day because they receive the same information. This lack of independence will inflate the test statistics, though it won't bias the observed proportions. For Hypotheses 1 and 2 the null hypotheses are rejected with such a high degree of statistical significance that some lack of independence is not problematic. In the following discussion, the data are, at times, divided into several partitions (e.g., Figure 2 and Table VI). Where t -statistics for individual partitions approach the conventional thresholds of statistical significance, they should be viewed with some skepticism.

To gain some perspective into how critical the independence assumptions made above are to the primary finding of this paper—that investors realize gains too soon and hold losers too long—it is instructive to look at an alternative test. Suppose that instead of assuming that independence exists at a transactional level we assume only that it exists at an account level. That is, we assume that the proportions of gains and losses realized in each account are independent of those realized in other accounts. PGR and PLR are then estimated for each account and their difference, PGR – PLR, is calculated for each account. The average account PGR is 0.57, the average account PLR is 0.36, the average of PGR – PLR is 0.21, and the hypothesis that the mean of PGR – PLR is less than or equal to zero is rejected with a t -statistic of 19.⁷ This alternative test also attempts to control for dependence caused by common information. To do this the sale of a stock is only counted if no sale has been previously counted for that stock in any account within a week before or after the sale date. That is, no two sales of the same stock within a week of each other are counted. Similarly, no two unrealized paper losses or gains of the same stock within a week of each other are counted. This test provides an alternative to the one reported in Table I and throughout the rest of the paper, but it is not without drawbacks. The previous test, in

⁶ To calculate the t -statistics in Table I, the standard error for the difference in the proportions PGR and PLR is:

$$\sqrt{\frac{\text{PGR}(1 - \text{PGR})}{n_{rg} + n_{pg}} + \frac{\text{PLR}(1 - \text{PLR})}{n_{rl} + n_{pl}}}$$

where n_{rg} , n_{pg} , n_{rl} , and n_{pl} are the number of realized gains, paper gains, realized losses, and paper losses.

⁷ An account is included in this test only if the denominators for both PGR and for PLR are nonzero for that account. There are 1893 such accounts. These same accounts are used to calculate share-based and dollar-based PGR and PLR.

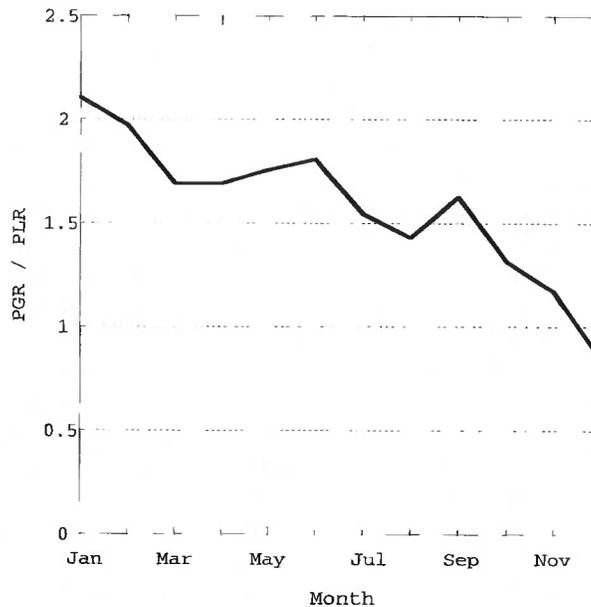


Figure 2. Ratio of the Proportion of Gains Realized (PGR) to the Proportion of Losses Realized (PLR) for each month. PGR is the number of realized gains divided by the number of realized gains plus the number of paper (unrealized) gains, and PLR is the number of realized losses divided by the number of realized losses plus the number of paper (unrealized) losses. Realized gains, paper gains, losses, and paper losses are aggregated over time (1987–1993) and across all accounts in the data set.

effect, weights each account by the number of realized and paper gains and losses in that account. This alternative test weights each account equally, which means we ignore the fact that accounts with more transactions provide more accurate estimates of their actual PGR and PLR. In other words, by treating each account the same, we assume that the observed account PGRs and PLRs are homoskedastic when they are clearly heteroskedastic. However, to properly weight for this heteroskedasticity we need to know the degree of independence of transactions within accounts, which is exactly the issue this test is intended to circumvent. It is presented here simply to demonstrate that when a different set of independence assumptions is made, the null to Hypothesis 1 is still rejected at a very significant level.

It should be noted that the PGR and the PLR measures are dependent on the average size of the portfolios from which they are calculated. When the portfolio sizes are large, both of these proportions will be smaller. Thus these proportions are smaller for traders who trade frequently and generally have larger portfolios than for those who trade less frequently. When PGR and PLR are calculated for Table I, the accounts with more trades weigh more heavily than those with fewer trades. In the alternative specification described in the last paragraph all accounts are weighted equally. For this

reason PGR and PLR are both larger in the alternative specification than in Table I. Of primary interest is not the individual values of PGR and PLR, but their values relative to each other.

Throughout this paper PGR and PLR are calculated in terms of trades and potential trades. An alternative specification is to calculate them in terms of number of shares traded and potential number of shares traded. When gains, losses, paper gains, and paper losses are aggregated across accounts before calculating PGR and PLR, as is done throughout most of this paper, measuring in shares further complicates the question of independence. However, if PGR and PLR are first calculated for each account and then the mean of $\text{PGR} - \text{PLR}$ is calculated—as in the alternative test described in the previous paragraph—independence is assumed only between accounts. When this alternative test is done for PGR and PLR based on shares rather than trades, the results are virtually unchanged: average PGR is 0.58, average PLR is 0.36, and the null hypothesis that the mean of $\text{PGR} - \text{PLR}$ is less than or equal to zero is rejected with a t -statistic of 18.⁸

Suppose investors frequently realize small gains and less frequently take large losses. It is then possible that they are selling similar proportions of the values of their gains and losses, though realizing gains at a higher rate on a trade-counted basis. This is, however, not the case. I calculate the average PGR and PLR per account by measuring losses, gains, potential losses, and potential gains in terms of dollars rather than shares or trades. When, as before, no two sales or potential sales of the same stock within a week of each other are counted, the average dollar-based PGR is 0.58 and the average dollar-based PLR is 0.42. The hypothesis that the mean of $\text{PGR} - \text{PLR}$ is less than or equal to zero is rejected with a t -statistic of 13.⁹

In Table I the ratio of PGR to PLR for the entire year is a little over 1.5, indicating that a stock that is up in value is more than 50 percent more likely to be sold from day to day than a stock that is down. In Weber and Camerer's (1995) experimental studies of the disposition effect, a stock that is up is also about 50 percent more likely to be sold than one that is down. Figure 2 charts the ratio of PGR to PLR for each month. This ratio declines from 2.1 in January to 0.85 in December. This decline is consistent with Constantinides' tax-loss selling model and suggests that at least some investors pay attention to tax-motivated selling throughout the year. From January through November, however, the observed ratio of PGR to PLR is greater

⁸ As in the previous test no two sales or potential sales of the same stock within a week of each other are counted here. If sales and potential within a week of each other are also counted, share-based PGR is 0.51 and share-based PLR is 0.31.

⁹ To examine this issue from another perspective, I look at each year in each portfolio and tally the total number of years for which both potential gains and potential losses are present in the portfolio and either net gains or net losses are realized. Net dollar gains are realized for 2,116 of these years and net dollar losses for 1,477 years. This indicates that, in most cases, large losses are not offsetting small gains.

than 1 and the hypothesis that the population ratio is less than or equal to 1 is rejected in each of these months with t -statistics ranging from 3.6 in November to 18 in January.¹⁰

To test the robustness of these results the data set is partitioned into two time periods and also into two groups of traders. Table II displays results when the data set is partitioned into stocks sold from 1987 to 1990 and 1990 to 1993, and when it is partitioned into the decile of traders who trade most frequently and the nine deciles of traders who trade least frequently. In the data set, the most active 10 percent of the traders account for 57 percent of all stock trades. In both time periods and for both the frequent and the infrequent traders, a significantly greater proportion of all possible gains than of all possible losses is realized throughout the year (t greater than 22, in all cases). In December, losses are realized at a higher rate relative to gains than during the rest of the year, indicating that investors are realizing tax losses in December. Due to how portfolios are reconstructed over time, average portfolio sizes are larger for the later years of the sample. PGR and PLR are therefore smaller for the second temporal partition, just as they are smaller for the partition of frequent traders.

One reason investors might choose to sell winners rather than losers is that they anticipate a change in the tax law under which capital gains rates will rise. The tax law of 1986 made such a change. If investors sold off winners in anticipation of higher tax rates, they might have entered 1987 with a larger percentage of losers in their portfolio than usual. Because such stocks are purchased prior to 1987 they would not show up in the portfolios reconstructed here. It is possible therefore that the rate at which winners are being realized relative to losers is lower in the investors' total portfolio than in the partial reconstructed portfolios. As old stocks are sold and new ones purchased the partial portfolios become more and more representative of the total portfolio. We would expect that if a sell-off of winners in anticipation of the 1986 tax law affects the observed rate at which gains and losses are realized in the partial portfolios, that effect would be greater in the first part of the sample period than in the last part. However the ratio PGR/PLR is virtually the same for the periods 1987 to 1990 and 1991 to 1993.

¹⁰ In Tables I-VI, realized and unrealized losses are tabulated on days that sales took place in portfolios of two or more stocks. One objection to this formulation is that for portfolios that hold only winners or only losers an investor cannot choose whether to sell a winner or to sell a loser, but only which winner or loser to sell. Another objection is that if an investor has net capital losses of more than \$3,000 for the current year (in non-tax-deferred accounts) it may be normative for that investor to choose to sell a winner rather than a loser. I have repeated the analyses reported in the tables subject to the additional constraints that there be at least one winner and one loser in a portfolio on the day of a sale for that day to be counted and that the net realized capital losses for the year to date in the portfolio be less than \$3,000. When these constraints are imposed, the difference in PGR and PLR is, for each analysis, greater. For example, for the entire sample and the entire year (as in Table I) there are 10,111 realized gains, 71,817 paper gains, 5,977 realized losses, and 94,419 paper losses. Thus the PLR is 0.060; the PGR is 0.123; their difference is 0.063; and the t -statistic for the difference in proportions is 47.

Table II
PGR and PLR Partitioned by Period and Trading Activity

This table compares the aggregate Proportion of Gains Realized (PGR) to the aggregate Proportion of Losses Realized (PLR), where PGR is the number of realized gains divided by the number of realized gains plus the number of paper (unrealized) gains, and PLR is the number of realized losses divided by the number of realized losses plus the number of paper (unrealized) losses. The data are partitioned into the periods 1987–1990 and 1991–1993 and into the 10 percent of the accounts that trade most frequently and the 90 percent that trade least frequently. For 1987–1990 there are 7,280 realized gains, 28,998 paper gains, 7,253 realized losses, and 50,540 paper losses. For 1991–1993 there are 6,603 realized gains, 50,660 paper gains, 4,677 realized losses, and 59,808 paper losses. For frequent traders there are 10,186 realized gains, 75,182 paper gains, 8,886 realized losses, and 103,096 paper. For infrequent traders there are 3,697 realized gains, 4,476 paper gains, 3,042 realized losses, and 7,251 paper losses. The *t*-statistics test the null hypotheses that the differences in proportions are equal to zero assuming that all realized gains, paper gains, realized losses, and paper losses result from independent decisions.

	1987–1990	1991–1993	Frequent Traders	Infrequent Traders
Entire year PLR	0.126	0.072	0.079	0.296
Entire year PGR	0.201	0.115	0.119	0.452
Difference in proportions	-0.075	-0.043	-0.040	-0.156
<i>t</i> -statistic	-30	-25	-29	-22
December PLR	0.143	0.110	0.095	0.379
December PGR	0.129	0.097	0.084	0.309
Difference in proportions	0.014	0.013	0.010	0.070
<i>t</i> -statistic	1.9	2.3	2.3	3.5
Jan.–Nov. PLR	0.123	0.069	0.078	0.282
Jan.–Nov. PGR	0.207	0.117	0.123	0.469
Difference in proportions	-0.084	-0.048	-0.045	-0.187
<i>t</i> -statistic	-32	-27	-31	-25

Table III reports the average returns since the day of purchase for realized and paper winners and losers. In December the losses that are realized are of much greater magnitude than those realized throughout the rest of the year. This is additional evidence that some investors do engage in tax-motivated selling in December.

Lakonishok and Smidt (1986) suggest that investors might sell winners and hold on to losers in an effort to rebalance their portfolios. We expect that investors who are rebalancing will sell a portion, but not all, of their shares of winning stocks. A sale of the entire holding of a stock is most likely not motivated by the desire to rebalance. So to eliminate trades that may be motivated by a desire to rebalance, I calculate PGR and PLR using only sales that are of an account's entire position in a stock (and using paper gains and losses on the days of those sales). There may be some cases where shares of a stock are already in the portfolio before 1987 and then additional shares are purchased. For these, the sale of all shares purchased after 1987

Table III
Average Returns

This table reports the mean return realized on stocks sold for a gain and on stocks sold for a loss. It also reports mean return that could be realized by stocks that are not sold on days that other stocks in the same portfolio are sold. These stocks are classified as paper gains and paper losses. For all accounts over the entire year, there are 13,883 realized gains, 79,658 paper gains, 11,930 realized losses, and 110,348 paper losses. For all accounts during the month of December, there are 866 realized gains, 7,131 paper gains, 1,555 realized losses, and 10,604 paper losses.

	Jan.-Nov.	December	Entire Year
Return on realized gains	0.275	0.316	0.277
Return on paper gains	0.463	0.500	0.466
Return on realized losses	-0.208	-0.366	-0.228
Return on paper losses	-0.391	-0.417	-0.393

may not amount to the sale of all shares held. So this removal of sales that could be motivated by diversification is not perfect. Even so, if the preference for selling winners is due to rebalancing, removing most rebalancing-motivated trades will greatly reduce the preference for selling winners.

In Table IV, for the entire year, when partial sales are ignored the preference for selling winners rather than losers is not substantially changed. The tendency to sell winners and hold losers does not appear to be the result of rebalancing. When partial sales are ignored, investors realize losses in December at an even higher rate relative to realizing gains. Perhaps this is because investors who are intentionally realizing tax losses choose to sell their entire position in the losing stock.

Investors who sell winners for the purpose of rebalancing their portfolios are likely to make new purchases. In an alternative effort to eliminate trades that may be motivated by a desire to rebalance, I calculate PGR and PLR using only sales for which there is no new purchase into a portfolio on the sale date or during the following three weeks (and using paper gains and losses on the days of those sales). Table V reports that when sales motivated by a desire to rebalance are eliminated in this way, investors continue to prefer to sell winners. Once again, investors realize losses at a higher rate than gains in December.

Another reason investors might sell winners and hold losers is that they expect the losers to outperform the winners in the future. An investor who buys a stock because of favorable information may sell that stock when it goes up because she believes her information is now reflected in the price. On the other hand, if the stock goes down she may continue to hold it, believing that the market has not yet come to appreciate her information. Investors could also choose to sell winners and hold losers simply because they believe prices mean revert. I test whether such beliefs are justified, *ex post*.

Table IV

PGR and PLR When the Entire Position in a Stock Is Sold

This table compares the aggregate Proportion of Gains Realized (PGR) to the aggregate Proportion of Losses Realized (PLR), where PGR is the number of realized gains divided by the number of realized gains plus the number of paper (unrealized) gains, and PLR is the number of realized losses divided by the number of realized losses plus the number of paper (unrealized) losses. In this table losses and gains are counted only if a portfolio's total position in a stock was sold that day. Paper (unrealized) gains and losses are counted only if the portfolio's total position in another stock held in the portfolio was sold that day. Realized gains, paper gains, losses, and paper losses are aggregated over time (1987–1993) and across all accounts in the dataset. PGR and PLR are reported for the entire year and for December only. For the entire year there are 10,967 realized gains, 36,033 paper gains, 9,476 realized losses, and 51,502 paper losses. For December there are 666 realized gains, 3,440 paper gains, 1,171 realized losses, and 4,759 paper losses. The *t*-statistics test the null hypotheses that the differences in proportions are equal to zero assuming that all realized gains, paper gains, realized losses, and paper losses result from independent decisions.

	Entire Year	December
PLR	0.155	0.197
PGR	0.233	0.162
Difference in proportions	-0.078	0.035
<i>t</i> -statistic	-32	4.6

Table VI reports excess returns for periods following the sale of a winning stock or the observation of a paper loss. Three investment horizons are examined: 84 trading days, which is the approximate median in sample holding period for stocks,¹¹ 254 trading days (one year), which is Benartzi and Thaler's (1995) estimate of the average investor's investment horizon, and 504 trading days (two years), which is how often, on average, New York Stock Exchange equities turned over during this period. Returns are calculated in excess of the CRSP value-weighted index. For winners that are sold, the average excess return over the following year is 3.4 percent more than it is for losers that are not sold. Investors who sell winners and hold losers because they expect the losers to outperform the winners in the future are, on average, mistaken. The superior returns to former winners noted here are consistent with Jegadeesh and Titman's (1993) finding of price momentum in security returns at horizons of up to eighteen months, though DeBondt and Thaler (1985, 1987) find price reversals at longer horizons of three to five years.¹²

¹¹ Note that the in-sample median holding period is a downwardly biased estimate of the true median holding period since stocks held for long periods are more likely to be bought before or sold after the data period and therefore not counted in the sample averages. The average turnover rate for equity in these accounts is 6.5 percent per month, which corresponds to an average holding period of about 15 months.

¹² At the time of this study CRSP data were available through 1994. For this reason two-year subsequent returns are not calculated for sales dates in 1993.

Table V
PGR and PLR When No New Stock Is Purchased
Within Three Weeks of Sale

This table compares the aggregate Proportion of Gains Realized (PGR) to the aggregate Proportion of Losses Realized (PLR), where PGR is the number of realized gains divided by the number of realized gains plus the number of paper (unrealized) gains, and PLR is the number of realized losses divided by the number of realized losses plus the number of paper (unrealized) losses. In this table losses and gains are counted only if a no new purchase was made into a portfolio on the day of the sale or within three weeks following the sale. Paper (unrealized) gains and losses are counted for days on which qualifying sales were made. Realized gains, paper gains, losses, and paper losses are aggregated over time (1987–1993) and across all accounts in the dataset. PGR and PLR are reported for the entire year and for December only. For the entire year there are 8,336 realized gains, 10,240 paper gains, 7,553 realized losses, and 19,370 paper losses. For December there are 590 realized gains, 1,024 paper gains, 1,194 realized losses, and 1,863 paper losses. The *t*-statistics test the null hypotheses that the differences in proportions are equal to zero assuming that all realized gains, paper gains, realized losses, and paper losses result from independent decisions.

	Entire Year	December
PLR	0.281	0.391
PGR	0.449	0.366
Difference in proportions	-0.168	0.015
<i>t</i> -statistic	-36	1.6

The average excess returns to winners sold in Table VI are determined by calculating excess buy-and-hold returns over the periods subsequent to each profitable sale of each stock and then taking an average that weighs each observation equally. Many stocks are sold for a profit on more than one date; sometimes the same stock is sold for a profit on the same date by more than one investor. Each of these sales is counted as a separate observation. The same procedure applies to paper losses. The *p*-values in Table VI are estimated by bootstrapping an empirical distribution for the difference in average excess buy-and-hold returns to realized winners and paper losses. This empirical distribution is generated under the null hypothesis that subsequent excess returns to realized winners and paper losers are drawn from the same underlying distribution. The methodology is similar to that of Brock, Lakonishok, and LeBaron (1992) and Ikenberry, Lakonishok, and Vermaelen (1995). Lyon, Barber, and Tsai (1998) test the acceptance and rejection rates for this methodology and find that it performs well in random samples. For each stock in the sample for which CRSP return data are available, a replacement stock is drawn (with replacement) from the set of all CRSP stocks of the same size decile and same book-to-market quintile as the original stock. Using the replacement stocks together with the original observation dates, average excess buy-and-hold returns are calculated for the 84, 252, and 502 trading days following the dates on which sales for a profit or paper losses are observed. These averages, and their differences, constitute one observation from the empirical distribution. One thousand such observa-

Table VI
Ex Post Returns

This table compares average returns in excess of the CRSP value-weighted index to stocks that are sold for a profit (winning stocks sold) and to stocks that could be, but are not, sold for a loss (paper losses). Returns are measured over the 84, 252, and 504 trading days subsequent to the sale of a realized winner and subsequent to days on which sales of other stocks take place in the portfolio of a paper loser. p -values refer to the frequency with which differences in excess returns over the same periods in the empirical (bootstrapped) distributions exceed the difference in excess returns observed in the data.

	Performance over Next 84 Trading Days	Performance over Next 252 Trading Days	Performance over Next 504 Trading Days
Average excess return on winning stocks sold	0.0047	0.0235	0.0645
Average excess return on paper losses	-0.0056	-0.0106	0.0287
Difference in excess returns (p -values)	0.0103 (0.002)	0.0341 (0.001)	0.0358 (0.014)

tions are made. The null hypothesis is rejected at the α percent level if the average subsequent excess return to realized winners minus that to paper losers in the data set is greater than the $(1 - \alpha)$ percentile average excess return to realized winners minus that to paper losers observed in the empirical distribution.

We saw in Table III (column 3) that investors are more likely to realize smaller, rather than larger, gains and losses. It may be that, due to regret aversion, investors are most loath to realize their greatest losses, and, due to tax consequences, they postpone realizing their greatest gains. Lower price ranges are likely to have a greater proportion of large losers and a smaller proportion of large winners than upper price ranges. Investors will therefore have a greater propensity to not sell losers in lower price ranges, and to not sell winners in higher price ranges.

Harris (1988) suggests that investors' reticence to sell losers may be due to their sensitivity to higher trading costs at lower stock prices. Table VII reports PGR and PLR for different price ranges and return ranges for January through November. Stocks with a price less than or equal to \$10, with prices greater than \$10 and less than or equal to \$25, and with prices higher than \$25, represent, respectively, 36 percent, 35 percent, and 29 percent of the data set. Partitioning on magnitude of return controls for the disproportionate numbers of large losers in the lower price range and of large winners in the top price range. The ranges for absolute value of return are: 0 to 0.15, 0.15 to 0.30, 0.30 to 0.50, and greater than 0.50. We see that in fourteen of fifteen partitions winners are realized at a higher rate than losers. This difference is statistically significant in thirteen partitions. When comparing winners and losers of similar magnitude, investors appear to prefer to sell winners and hold losers even when trading costs for both are about the same.

Table VII
PGR and PLR Partitioned by Price and Return

This table compares the aggregate Proportion of Gains Realized (PGR) to the aggregate Proportion of Losses Realized (PLR), where PGR is the number of realized gains divided by the number of realized gains plus the number of paper (unrealized) gains, and PLR is the number of realized losses divided by the number of realized losses plus the number of paper (unrealized) losses. The data are partitioned on stock price and on absolute value of the return to date (R), for all accounts, 1987–1993, January through November only. The t -statistics test the null hypotheses that the differences in proportions are equal to zero assuming that all realized gains, paper gains, realized losses, and paper losses result from independent decisions.

	$ R \leq 0.15$	$0.15 < R \leq 0.30$	$0.30 < R \leq 0.50$	$0.50 \leq R $
Price \leq \$10				
PLR	0.141	0.129	0.109	0.030
PGR	0.267	0.257	0.295	0.282
Difference	-0.126	-0.128	-0.186	-0.252
t -statistic	13.0	10.2	11.7	17.5
\$10 \leq Price \leq \$25				
PLR	0.138	0.105	0.076	0.058
PGR	0.222	0.186	0.172	0.135
Difference	-0.084	-0.081	-0.096	-0.077
t -statistic	16.9	13.1	13.1	11.3
\$25 \leq Price				
PLR	0.125	0.104	0.104	0.049
PGR	0.197	0.126	0.081	0.055
Difference	-0.072	-0.022	0.023	-0.006
t -statistic	19.4	4.5	-3.2	0.61

There is another way to contrast the hypothesis that losses are realized more slowly due to the higher transactions costs with the two behavioral hypotheses. We can look at the rates at which investors purchase additional shares of stocks they already own. The proportion of gains purchased again (PGPA) and the proportion of losses purchased again (PLPA) can be calculated in a manner analogous to how PGR and PLR are calculated. When a stock already in the portfolio is purchased again it is counted as a gain purchased again or a loss purchased again. On days when purchases are made, stocks already in the portfolio for which additional shares are not repurchased are counted as gains or losses potentially purchased again. Thus:

$$\frac{\text{Gains Purchased Again}}{\text{Gains Purchased Again} + \text{Gains Potentially Purchased Again}} = \frac{\text{Proportion of Gains Purchased Again}}{\text{Proportion of Gains Purchased Again}} \text{ (PGPA)} \quad (3)$$

$$\frac{\text{Losses Purchased Again}}{\text{Losses Purchased Again} + \text{Losses Potentially Purchased Again}} = \frac{\text{Proportion of Losses Purchased Again}}{\text{Proportion of Losses Purchased Again}} \text{ (PGPA)} \quad (4)$$

When these proportions are calculated, additional purchases of a particular stock in a particular account are not counted if they take place within one week of a previous purchase of the stock. This is done to avoid the possibility of counting a purchase order filled over more than one day as an additional purchase.

If investors avoid the higher transactions cost of low priced stocks we would expect PLPA to be less than PGPA. If, however, investors are more risk seeking for losing investments (prospect theory) or if they believe prices will revert (as do Andreassen's subjects), then PLPA will be greater than PGPA. This is the case. For the entire sample $PLPA = 0.135$ and $PGPA = 0.094$. If we assume that all decisions to purchase or not purchase additional stock are independent, the hypothesis that these two proportions are equal can be rejected with a t -statistic of 19. This supports the two behavioral theories, but not the transaction cost hypothesis.¹³

In Table III we saw that investors tend to sell their larger gains and losses at a slower rate than their smaller gains and losses. Prospect theory does not predict that investors realize their large gains more slowly than their small gains. Nor does a belief in mean reversion predict this. If, however, investors believe that stocks that perform moderately well will revert, but those that perform unusually well will trend,¹⁴ they might sell their small winners and hold their larger ones. These beliefs could then also lead them to buy fewer additional shares of small winners and more additional shares of larger winners. To test this I partition winning investments into large or small winners using the mean unrealized winners' return of 0.47 as a break point (see Table III). Similarly I partition losers into large and small losers using -0.39 as a break point. Small winners are repurchased at a rate of 0.112, large winners are repurchased at a rate of 0.043. Small losers are repurchased at a rate of 0.172 and large losers at 0.067. The difference in the rates at which large and small gains are realized is highly significant (t equals 26, assuming independence); so, too, is the difference in the rate at which large and small losses are realized (t equals 39). These investors do not tend to buy additional shares of big winners. This is not consistent with the hypothesis that they believe small winners will revert but large winners will perform well, however other factors may be working against the hypothesis. Investors who are in the habit of buying additional shares of stocks they already own may reach their limit of additional purchases before these stocks have an opportunity to make large gains (or losses). Regret aversion may also influence investors to not buy additional shares of big winners. For example, suppose an investor buys 100 shares of stock A at \$100 per share. Then stock A appreciates to \$150. The investor may believe stock A will

¹³ For the same reasons as discussed in Section III B, these decisions will not always be independent. So the t -statistic of 19 overstates the actual statistical significance.

¹⁴ In this vein Barberis, Shleifer, and Vishny (1996) develop a model in which investors believe that earnings switch between two regimes, one mean reverting and the other trend following.

continue to appreciate but he may still refrain from buying an additional 100 shares, for if he does purchase more shares, he will more poignantly regret that he didn't buy them at \$100 per share to begin with. The greater the difference between the original and additional purchase prices, the greater is this potential regret.

The results presented so far are not able to distinguish between the two behavioral hypotheses. Both prospect theory and a belief in mean reversion predict that investors will hold their losers too long and sell their winners too soon. Both predict that investors will purchase more additional shares of losers than of winners. However a belief in mean reversion should apply to stocks that an investor does not already own as well as those she does, but prospect theory applies only to the stocks she owns. Thus a belief in mean reversion implies that investors will tend to buy stocks that had previously declined even if they don't already own these stocks, and prospect theory makes no prediction in this case. Odean (1997) finds that this same group of investors tends to buy stocks that have, on average, outperformed the CRSP value-weighted index by about 25 percent over the previous two years. This would appear inconsistent with a simple belief in mean reversion. (It is, though, consistent with a belief that big winners will continue to perform well.)

IV. Discussion

This paper examines the behavior of individual investors and finds that investors exhibit disposition effects; that is, they realize their profitable stocks investments at a much higher rate than their unprofitable ones, except in December. The extent to which this behavior affects market prices depends on the trading activities of other market participants such as professional traders and institutional investors. If the disposition effect holds in aggregate it may contribute to the positive relationship between price change and volume identified by Lakonishok and Smidt (1986) and by Ferris et al. (1988). The disposition effect could also be a cause of the positive correlation between price changes and volume in other markets such as residential real estate. Case and Shiller (1988) report evidence of disposition effects from interviews with homeowners in boom and post-boom real estate markets.

By affecting supply, the disposition effect may also contribute to market stability near prices at which substantial trading has previously taken place. If many investors buy a stock at a particular price, that price may become their reference point. If the stock falls below this reference point, these investors will be averse to selling for a loss, reducing the supply of potential sellers. A reduced supply of potential sellers could slow further price decreases. On the other hand, if the stock rises above the reference point, these investors will be more willing to sell, increasing the supply of potential sellers, and possibly slowing further price increases. If these investors have private information about the future prospects of a company whose stock

they hold, the disposition effect may slow the rate at which this information is incorporated into price. For example, investors with negative information may be unwilling to sell a stock if its price is below their reference point. In not selling the stock, these investors will fail to signal their negative information to the market, and there could be a delay before that information is reflected in prices.

Though the disposition effect may influence market prices, its economic significance is likely to be greatest for individual investors. To get a rough idea of the economic costs of the loss aversion, let us imagine that a hypothetical investor is choosing to sell one of two stocks. The first of these stocks behaves like the average realized winner in this data set and the other like the average paper loser. The investor wishes to sell \$1,000 worth of stock after commissions and that happens to be what his position in each stock is currently worth. Suppose he is averse to realizing losses and so sells the winning stock. If his experience is similar to that of the average investor in this data set, his return on the sale will be 0.277 (Table III, third column). Since the stock is currently worth \$1000, its purchase price must have been \$783, and his capital gain is \$217. If he instead chooses to sell \$1,000 worth of the losing stock, his return will be -0.393 , with a purchase price of \$1,647, and a capital loss of \$647. One year later the (losing) stock that he held will have, on average, a return 1 percent below the market (Table VI); the winning stock that he sold will have, on average, a return 2.4 percent above the market. Marginal tax rates for capital gains for investors in this sample vary from 0 to 28 percent, plus state taxes. Assume that our investor's marginal tax rate is 15 percent and that he has taxable gains against which to offset losses. Then by choosing to sell the winning stock rather than the loser, he gives up an immediate tax savings of \$130. Suppose that whichever stock the investor does not sell now, he will sell in one year; then the investor is paying \$130 in taxes one year earlier than he otherwise would. If he can expect a return of 8 percent on his money, choosing not to defer these taxes costs him about \$10. In addition to this, over the next year the investor's return on the stock he holds (the loser) is \$34 less than if he had held the other stock (the winner). Using \$1,000 as a basis, and including the value of the immediate tax savings (\$10) as well as the anticipated difference in capital gains (\$34), the investor's return is about 4.4 percent higher over the next year if he sells the loser rather than the winner. The benefits of deferring taxes may be even higher if the investor chooses to delay realizing his gains for more than one year. On the other hand, a habit of regular loss realizations may reduce the magnitude of available capital losses.

The trading records analyzed in this paper are obtained from a discount brokerage house. This avoids the need to consider agency issues that influence institutional investors or to disentangle the decisions and motivations of individual investors from those of their retail brokers. It would be illuminating to repeat this study with data on institutional trading and with data from a retail brokerage house.

V. Conclusion

This paper finds that individual investors demonstrate a significant preference for selling winners and holding losers, except in December when tax-motivated selling prevails. This investor behavior does not appear to be motivated by a desire to rebalance portfolios or by a reluctance to incur the higher trading costs of low priced stocks. Nor is it justified by subsequent portfolio performance. It leads, in fact, to lower returns, particularly so for taxable accounts.

REFERENCES

- Andreassen, Paul, 1987, On the social psychology of the stock market: Aggregate attributional effects and the regressiveness of prediction, *Journal of Personality and Social Psychology* 53, 490-496.
- Andreassen, Paul, 1988, Explaining the price-volume relationship: The difference between price changes and changing prices, *Organizational Behavior and Human Decision Processes* 41, 371-389.
- Andreassen, Paul, 1990, Judgmental extrapolation and market overreaction: On the use and disuse of news, *Journal of Behavioral Decision Making* 3, 153-174.
- Badrinath, S., and Wilber Lewellen, 1991, Evidence on tax-motivated securities trading behavior, *Journal of Finance* 46, 369-382.
- Barberis, Nicholas, Andre Shleifer, and Robert Vishny, 1996, A model of investor sentiment with both underreaction and overreaction, Working paper, University of Chicago.
- Benartzi, Shlomo, and Richard Thaler, 1995, Myopic loss aversion and the equity premium puzzle, *Quarterly Journal of Economics* 110, 73-92.
- Bremer, Marc, and Kato Kiyoshi, 1996, Trading volume for winners and losers on the Tokyo Exchange, *Journal of Financial and Quantitative Analysis* 31, 127-142.
- Brock, William, Josef Lakonishok, and Blake LeBaron, 1992, Simple technical trading rules and the stochastic properties of stock returns, *Journal of Finance* 47, 1731-1764.
- Case, Karl, and Robert Shiller, 1988, The behavior of home buyers in boom and post-boom markets, *New England Economic Review* November/December, 29-46.
- Constantinides, George, 1984, Optimal stock trading with personal taxes: Implications for prices and the abnormal January returns, *Journal of Financial Economics* 13, 65-69.
- DeBondt, Werner, and Richard Thaler, 1985, Does the stock market overreact?, *Journal of Finance* 40, 793-807.
- DeBondt, Werner, and Richard Thaler, 1987, Further evidence on investor overreaction and stock market seasonality, *Journal of Finance* 42, 557-581.
- Dyl, Edward, 1977, Capital gains taxation and the year-end stock market behavior, *Journal of Finance* 32, 165-175.
- Ferris, Stephen, Robert Haugen, and Anil Makhija, 1988, Predicting contemporary volume with historic volume at differential price levels: Evidence supporting the disposition effect, *Journal of Finance* 43, 677-697.
- Harris, Lawrence, 1988, Discussion of predicting contemporary volume with historic volume at differential price levels: Evidence supporting the disposition effect, *Journal of Finance* 43, 698-699.
- Heisler, Jeffrey, 1994, Loss aversion in a futures market: An empirical test, *Review of Futures Markets* 13, 793-822.
- Ikenberry, David, Josef Lakonishok, and Theo Vermaelen, 1995, Market underreaction to open market share repurchases, *Journal of Financial Economics* 39, 181-208.
- Jegadeesh, Narasimhan, and Sheridan Titman, 1993, Returns to buying winners and selling losers: Implications for stock market efficiency, *Journal of Finance* 48, 65-91.

- Kahneman, Daniel, and Amos Tversky, 1979, Prospect theory: An analysis of decision under risk, *Econometrica* 46, 171-185.
- Lakonishok, Josef, and Seymour Smidt, 1986, Volume for winners and losers: Taxation and other motives for stock trading, *Journal of Finance* 41, 951-974.
- Lyon, John, Brad Barber, and Chih-Ling Tsai, 1998, Improved methods for tests of long-run abnormal stock returns, *Journal of Finance*, forthcoming.
- New York Stock Exchange Fact Book*, 1995, (New York Stock Exchange, Inc., New York, N.Y.).
- Odean, Terrance, 1997, Do investors trade too much?, Working paper, University of California, Davis.
- Poterba, James, 1987, How burdensome are capital gains taxes? Evidence from the United States, *Journal of Public Economics* 33, 157-172.
- Schlarbaum, Gary, Wilber Lewellen, and Ronald Lease, 1978, Realized returns on common stock investments: The experience of individual investors, *Journal of Business* 51, 299-325.
- Shefrin, Hershey, and Meir Statman, 1985, The disposition to sell winners too early and ride losers too long: Theory and evidence, *Journal of Finance* 40, 777-790.
- Starr-McCluer, Martha, 1995, Tax losses and the stock portfolios of individual investors, Working paper, Federal Reserve Board of Governors.
- Thaler, Richard, 1985, Mental accounting and consumer choice, *Marketing Science* 4, 199-214.
- Weber, Martin, and Colin Camerer, 1998, The disposition effect in securities trading: An experimental analysis, forthcoming *Journal of Economic Behavior and Organization*.

78

Review of Selected Studies and Comments in Response to the Department of Labor's Conflict of Interest 2015 Proposed Rule and Exemptions

March 4, 2016

Advanced Analytical Consulting Group, Inc.

Karthik Padmanabhan, MBA, MS 312-551-9001 karpad@aacg.com	Constantijn Panis, PhD 213-784-6400 stanpanis@aacg.com	Timothy J. Tardiff, PhD 617-827-4043 timtardiff@aacg.com
--	---	---

ABSTRACT

In April 2015, the U.S. Department of Labor (DOL) published a Conflict of Interest Proposed Rule.¹ The DOL received numerous comments. This document reviews six studies that were submitted among the comments by NERA Economic Consulting, Oliver Wyman, the Investment Company Institute, Compass Lexecon, Robert Litan and Hal Singer of Economists Inc., and Quantria Strategies.

We first discuss a number of common themes that were raised in the studies and then separately address each of the six studies with a summary of our opinions, a synopsis of the study, and a discussion of the pertinent arguments. We generally find the studies lacking in rigor, failing to recognize emerging alternatives to traditional offerings of investment advice, incorrectly equating the benefits of conflicted advice to those of non-conflicted advice, or suffering from logical fallacies. None of the studies offer compelling arguments against implementation of the DOL's Conflict of Interest Proposed Rule.

¹ See Federal Register, Volume 80, p. 21928, available at <http://webapps.dol.gov/FederalRegister/PdfDisplay.aspx?DocId=28201>.

CONTENTS

Abstract	i
Abbreviations	iii
1. Introduction	1
2. Common Themes	3
Common Theme 1: Investors with Higher Balances Will Migrate to Higher-Cost Fee-Based Accounts	3
Common Theme 2: Investors with Low Balances Will Lose Access to Advice.....	5
Common Theme 3: Lack of Advice Will Cause Investors to Make Mistakes and Save Less	8
Common Theme 4: The Academic Literature Is Misapplied	9
3. NERA Economic Consulting	12
Summary.....	12
Synopsis.....	13
Discussion.....	14
4. Oliver Wyman.....	22
Summary.....	22
Synopsis.....	23
Discussion.....	23
5. Investment Company Institute	32
Summary.....	32
Synopsis	32
Discussion.....	33
6. Compass Lexecon	41
Summary.....	41
Synopsis	42
Discussion.....	42
7. Litan and Singer	45
Summary.....	45
Synopsis	47
Discussion.....	47
8. Quantria Strategies	56
Summary.....	56
Synopsis	56
Discussion.....	57
9. References	64
Disclaimer	69

ABBREVIATIONS

bps	Basis points (1 bps = 0.01%)
DB	Defined benefit
DC	Defined contribution
DOL	U.S. Department of Labor
ERISA	Employee Retirement Income Security Act
EBRI	Employee Benefit Research Institute
GAO	General Accounting Office
ICI	Investment Company Institute
IRA	Individual Retirement Account (or Individual Retirement Arrangement)
RIA	Regulatory impact analysis
S&P	Standard & Poors
SCF	Survey of Consumer Finances
SEC	Securities and Exchange Commission
SIFMA	Securities Industry and Financial Markets Association

1. INTRODUCTION

In April 2015, the DOL published a Conflict of Interest Proposed Rule ("Proposed Rule")² along with a Regulatory Impact Analysis (DOL 2015).³ This document contains reviews of six studies and supplemental materials that were submitted among the many comments that the DOL received.

- *NERA Economic Consulting*: SIFMA submitted comments, which included a comment by NERA Economic Consulting (NERA 2015a).⁴ In response to questions from the DOL, NERA provided additional details in a memorandum (NERA 2015b).⁵ No individual authors are listed on NERA (2015a). The author of NERA (2015b) is Patrick Conroy.
- *Oliver Wyman*: Several financial services firms jointly commented based on a study by Oliver Wyman Inc. titled "The role of financial advisors in the US retirement market" (Oliver Wyman 2015).⁶ No individual authors are listed.
- *Investment Company Institute*: Brian Reid and David W. Blass of ICI filed a July 21, 2015 letter "re: RIN 1210-AB32: Regulatory Impact Analysis, Definition of the Term "Fiduciary"; Conflict of Interest Rule—Retirement Advice" (ICI 2015a).⁷ In response to questions from the DOL, Reid and Blass provided additional details in two letters of September 24, 2015 (ICI 2015b)⁸ and December 1, 2015 (ICI 2015c).
- *Compass Lexecon*: Counsel for Primerica, Inc. submitted a study by Compass Lexecon titled "Tax Consequences to Investors Resulting from Proposed Rules Relating to Financial Representative Fiduciary Status" (Compass Lexecon 2015).⁹ No individual authors are listed.
- *Litan and Singer*: The Capital Group submitted a study by Robert Litan and Hal Singer of Economists Inc. on "Good Intentions Gone Wrong: The Yet-To-Be Recognized Costs of the Department of Labor's Proposed Fiduciary Rule" (Litan and Singer 2015a).¹⁰ In response to questions from the DOL, Litan and Singer provided additional details in a letter of September 21, 2015 (Litan and Singer 2015b).¹¹
- *Quantria Strategies*: On behalf of a group of clients, Davis & Harman LLP submitted a study by Quantria Strategies LLC titled "Unintended Consequences: Potential of the DOL Regulations to Reduce Financial Advice

² See Federal Register, Volume 80, p. 21928, available at <http://webapps.dol.gov/FederalRegister/PdfDisplay.aspx?DocId=28201>.

³ Available at <http://www.dol.gov/ebsa/pdf/conflictsofinterestria.pdf>.

⁴ Available at <http://www.dol.gov/ebsa/pdf/1210-AB32-2-00506.pdf>.

⁵ Available at <http://www.dol.gov/ebsa/pdf/1210-AB32-2-03079.pdf>.

⁶ Available at <http://www.dol.gov/ebsa/pdf/1210-AB32-2-00515.pdf>.

⁷ Available at <http://www.dol.gov/ebsa/pdf/1210-AB32-2-00749.pdf>.

⁸ Available at <http://www.dol.gov/ebsa/pdf/1210-AB32-2-03056.pdf>.

⁹ Available at <http://www.dol.gov/ebsa/pdf/1210-AB32-2-00615.pdf>.

¹⁰ Available at <http://www.dol.gov/ebsa/pdf/1210-AB32-2-00517.pdf>.

¹¹ Available at <http://www.dol.gov/ebsa/pdf/1210-AB32-2-02967.pdf>.

and Erode Retirement Readiness” (Quantria 2015).¹² No individual authors are listed.

Several studies made the same or similar assertions or arguments. We address some of these common themes in Section 2. Sections 3 through 8 review the six studies, each with a summary of our opinions, a synopsis of the study, and a detailed discussion.

¹² Available at <http://www.dol.gov/ebsa/pdf/1210-AB32-2-00746.pdf>.

2. COMMON THEMES

Several studies reviewed in this document made similar assertions or arguments. This section discusses such common themes. They include:

1. The Proposed Rule would force IRA investors with higher balances to migrate to higher-cost fee-based accounts;
2. The Proposed Rule would force IRA investors with low balances to lose access to advice;
3. Lack of advice prompted by the Proposed Rule would cause investors to make mistakes and save less; and
4. The Regulatory Impact Analysis (RIA) misapplies the academic literature.

We discuss these themes in turn, but first define typical IRA account types. IRAs may be held in arrangements through which the account holder has access to financial advice, such as a brokerage account or an advisory account. A brokerage account charges commissions, which may include a fixed amount per trade, a front-end or back-end load charge, annual 12b-1 fees, et cetera. These amounts (and the shares that are paid to the broker) may differ across financial products, which may give rise to conflicts of interest for brokers. In contrast, an advisory account typically charges a percentage of assets under management irrespective of the financial products in which the account balance is invested. This annual fee is also known as a wrap fee. IRAs may also be held in arrangements without professional advice, such as in a discount brokerage account. Consistent with most commentators, we use the term "brokerage account" for accounts that include access to advice and "discount brokerage account" for accounts without access to advice.

Common Theme 1: Investors with Higher Balances Will Migrate to Higher-Cost Fee-Based Accounts

The first and second common themes follow from the premise that the Proposed Rule would make current commission-based brokerage accounts unworkable, forcing the closure of such accounts. To preserve access to advice, some IRAs would migrate to fee-based advisory accounts. According to the first common theme, since a financial institution's main source of revenue from advisory accounts is an annual percentage of assets, migration would be profitable for larger accounts only. Conversely, the second common theme posits that smaller accounts would migrate to a discount brokerage account and lose access to advice.

Several studies argue that the move to advisory accounts would imply higher costs for the IRA account holder. Examples include the following: "This suggests that investors would pay more if moved to fee-based accounts" (NERA 2015a, p. 6); "Almost all retail investors would face increased costs (73% to 196% on average) from providers shifting clients to a fee-based advisory model" (Oliver Wyman 2015, p. 38); "if the account is large enough, move to an advisory relationship, which may increase fees, especially for buy and hold investors" (Quantria 2015, p. 7).

These claims are not based on empirical evidence of investor responses to fee changes. Further, they do not accurately reflect the empirical evidence about the full cost to investors of brokerage accounts. Instead, incorrect cost estimates and a

flawed assumption of static prices and service levels result in biased cost comparisons. First, the difference in fees charged to investors in advisory accounts versus brokerage accounts is smaller than the studies purport to document. Second, brokerage account holders have opted for a lower average service level than holders of advisory accounts and may continue to be served at a lower level (and at a lower cost) after migration.

Oliver Wyman (2015), Litan and Singer (2015a), and Quantria (2015) cite a 2011 study by Oliver Wyman which tabulated higher fees for advisory than for brokerage accounts. However, that comparison accounted for direct expenses only. It ignored expenses that the account holder paid to a third party which shared the proceeds with the broker. In the words of Oliver Wyman (2011, p. 22), the comparison “[e]xcludes marketing and distribution, shareholder services, and other fees not directly paid by investors.” The excluded cost components are predominantly relevant for brokerage accounts, i.e., the comparison is biased to make brokerage accounts look less expensive. The excluded cost components can be substantial. For example, 12b-1 fees and shareholder service fees can run as high as 100 bps (SEC 2015).

NERA (2015a) analyzed a proprietary data set of about 63,000 IRAs in brokerage and advisory accounts. It, too, compared expenses of advisory accounts to those of brokerage accounts and concluded that advisory accounts were more expensive. And like Oliver Wyman (2011), NERA (2015a, p. 4) biased the comparison by considering only direct fees: “Fees exclude revenue that the firm may receive indirectly from the account-holder, such as markup/markdown revenue or 12b-1 fees.” These and other indirect revenue components vary across products, tend to constitute conflicted compensation, and their exclusion therefore makes brokerage accounts appear less expensive than they really are. In a follow-up memorandum prompted by questions from the DOL, NERA (2015b) defended the exclusion of indirect fees with the assertion that its data set did not contain information related to such fees. However, the detailed, account-level data that NERA compiled presumably included information on fund positions, and 12b-1 fees for individual funds are readily available from Morningstar and other sources. In other words, NERA’s fee comparison is biased, making brokerage investments appear to have lower costs than they actually do. NERA acknowledged the bias and did not do anything to mitigate it even though doing so would have been relatively straightforward with publicly available information.

While Oliver Wyman’s and NERA’s expense comparisons bias brokerage expenses downward, even if average expenses for advisory accounts were higher than for brokerage accounts, a simple comparison of average expenses in brokerage and advisory accounts would not support a conclusion on whether expenses in brokerage accounts would rise or fall when migrated to advisory accounts. The level of activity in brokerage accounts tends to be much lower than that in advisory accounts, and the level of service required to maintain those accounts is correspondingly lower. For example, NERA (2015a) reported that the median brokerage account in its data file traded 6 times in 2014, compared with 57 times for the median advisory account. Advisory accounts tend to have higher balances, which explains some of the difference, but NERA (2015a) found trading frequencies among brokerage accounts to be much lower than among advisory accounts at all reported balance ranges. NERA (2015a, p. 7) presented its results in terms of self-selection of investors: “Thus, the data are consistent with the idea that investors who expect to trade often rationally choose fee-based accounts whereas those that do not trade often are likely to choose commission-based accounts.” None of the studies we reviewed suggested

that trading patterns would change if, as asserted, brokerage accounts are converted into advisory accounts. In other words, financial institutions may be expected to continue to provide the same, relatively low level of service after conversion as they currently provide to brokerage accounts.

In a competitive market, a lower level of service should of course be provided at a lower cost. Indeed, as observed by Council of Economic Advisers (2015, p. 21):

The cost of advice depends primarily on the resources necessary to provide it—the adviser’s time, IT infrastructure, and other inputs—rather than the form of the adviser’s compensation. Thus, an adviser receiving payment through non-conflicted structures should be able to provide advice at the same cost as an adviser receiving conflicted payments, as long as the inputs in time and infrastructure are equal.

The studies reviewed in this document failed to recognize that services currently provided to brokerage accounts should cost roughly the same in advisory accounts. For example, ICI (2015a, p. A-1) assumed—without motivation—that current pricing models will carry over: “The difference in the fees [between advisory and brokerage accounts] is roughly 60 basis points [...], which is the additional amount that each investor moving to a fee-based account would pay.” Instead of retaining their current pricing structures, financial institutions may be expected to offer modified account types that avoid fees on services that newly migrated investors do not demand.

Common Theme 2: Investors with Low Balances Will Lose Access to Advice

As noted above, several studies adopted the premise that the Proposed Rule would make current commission-based brokerage accounts unworkable and that it would force the closure of such accounts. The studies argued that larger IRAs would migrate to fee-based advisory accounts, but that such accounts would not be profitable for smaller accounts. Instead, they argue, smaller accounts would migrate to an account type without access to advice. Examples include the following: “If the DOL proposal were to make commission-based accounts unworkable for broker-dealers, these accounts [with balances under \$25,000] could no longer be maintained” (NERA 2015a, p. 12); “Millions of existing small balance IRA owners are likely to lose access to the financial advisor of their choice or any financial advisor at all” (Oliver Wyman 2015, p. 3); “it is very likely that under the current proposal investors with less than \$100,000 in IRA balances would not be able to get access to fee-based accounts” (ICI 2015a, p. A-1); and “Faced with this new [fiduciary] duty for brokerage accounts, many brokerage firms would likely react either by exiting the segment of the IRA market represented by individuals with modestly sized portfolios [...] or by switching to a fee-based advisory model for these investors” (Litan and Singer 2015a, p. 12). The studies vary in their assessment of a balance threshold below which advisory accounts would be unprofitable. ICI (2015a) assumes that the minimum balance for an advisory account is \$100,000; others contend the minimum balance may be as low as \$25,000.

Many factors cast doubt on the studies’ premise that IRA account holders with low balances will lose access to advice.

First, smaller investors already have advisory accounts despite assertions that low-balance advisory accounts are not profitable. The data described by NERA (2015a,

2015b) show that approximately 8% of IRAs with balances under \$25,000 are advisory accounts, and 20% of IRAs under \$100,000 are advisory accounts (compared with 29% across all account sizes).¹³ Evidence from overseas similarly suggests that low-balance accounts can be serviced profitably in the absence of conflicts of interest. For example, among Dutch accounts whose advisers "are paid fixed wages only," Kramer (2012) found that the 5th percentile of account balances was just €600.

Second, the Proposed Rule contains carve-outs and exemptions that seek to preserve current business models. Citing unspecified sources or no sources at all, several studies argue that the exemptions are unworkable. However, it is ultimately an empirical question to what extent financial institutions will take advantage of available exemptions.

Third, in addition to the two points above which are enough to demonstrate that smaller investors can and do have advisory accounts, financial institutions may develop new account types, or adjust current-style brokerage and advisory accounts. The perspective that current-style brokerage and advisory accounts, with current-style pricing structures, will be the only options available to IRA investors after conflict-of-interest regulations go into effect is too static. The declining cost of providing advice and related services has already created opportunities for new account types. The marketplace for new advisory options is rapidly changing with the arrival of automated or 'robo' investment advice. The minimum balance requirement for many of these robo-advisers is low enough to cater to IRA accounts with assets under \$25,000. For example, Tergesen (2015) documented that many robo-advisory firms, including such well-known investment advisory firms as Wealthfront Inc. and Betterment LLC have minimum balance requirements ranging from \$500 to as low as \$0. By definition, lower-balance accounts have fewer assets to invest and likely require fewer services than larger accounts. This characteristic can make them particularly suitable for automated advice. Also, target date funds rebalance automatically and adjust their exposure to risk automatically, thereby reducing the effort required to maintain an account.

In addition to the existing options, new options are appearing in the marketplace. For example, investors who strongly prefer human-based investment advice may have alternative options. Innovation in the marketplace for investment advice includes the advent of a hybrid model that combines automated and human-based investment advice methods. FutureAdvisor and SigFig, two hybrid online investment management advisers, offer access to an investment adviser to accounts with a minimum balance of \$10,000 (FutureAdvisor 2015, SigFig 2016). Another hybrid investment adviser, Personal Capital, recently lowered its minimum required to open an account from \$100,000 to \$25,000 and is part of an industry trend towards lower minimum balance requirements.

¹³ NERA (2015b, p. 4) shows that about 9% of fee-based accounts and 41% of commission-based accounts had balances under \$25,000. NERA (2015a, p. 4) states that 29.36% of accounts were fee-based, i.e., $9\% \times 29.36\% = 2.6\%$ of accounts were fee-based with a balance under \$25,000 and $41\% \times (1-29.36\%) = 29.0\%$ were commission-based with a balance under \$25,000. Fee-based accounts thus make up about $2.6\% / (2.6\% + 29.0\%) = 8\%$ of all accounts with balances under \$25,000. Analogous logic shows that about 20% of accounts with balances under \$100,000 were fee-based. Also see Table 4 on page 40 of this document.

Major brokerage houses that currently offer commission-based IRA accounts are entering the automated advice market and making automated advice a viable alternative to traditional investment advice models. A recent study forecasts that robo-advisory services will manage more than \$2 trillion in assets in just four years from now (AT Kearney 2015).

Fourth, additional evidence from overseas suggests that concerns over the loss of advice failed to materialize because investors who stopped being served found a replacement adviser. In a preliminary evaluation of changes to the financial system in the United Kingdom, which recently banned payments to financial advisers that depend on the advice given, Europe Economics (2014, p. 63) found:

Some advisers have sought to terminate unprofitable client relationships. Data from NMG Consulting, for example, imply that in the year to Q1 2014 about 310,000 clients stopped being served for this reason. On the other hand 820,000 clients were gained in the same period. The same survey indicates that advisers refused to serve about 60,000 (potential new) clients in the same period. If we assume that many of those clients with relationships terminated on the grounds of inadequate profitability sought out another adviser, the positive net increase in customers served suggests that such looking around for a replacement was largely successful. We cannot rule out the existence of a residual group of consumers denied service in this way. However these data do not speak to a significant issue here.

Fifth, the argument that an investor can be served profitably in a brokerage account but not in an advisory account raises the uncomfortable question of how advisers are able to serve small brokerage accounts under the current regime. The cost of providing advice depends not on the adviser's compensation structure but on the adviser's time, IT infrastructure, and other inputs. Suppose an adviser requires, say, \$500 annually to serve a \$20,000 account. If conflicted payments were banned, she could charge an asset-based fee of 2.5% (or a flat \$500) and continue providing financial advice. The account holder may object to such charges as too high and decide to forego the advice. This scenario is consistent with the argument that advisers would be unable to charge fees sufficient to cover their costs. At issue is why the account holder and the adviser have a relationship under the current regime. A plausible explanation is that the account holder does not realize how much he is paying for advice; once confronted with the full cost, he is not willing to buy it anymore. This implies a market failure: brokerage account holders may currently purchase too much conflicted advice. Basic economics suggest that a reduction of advice, resulting from greater transparency of costs and fees, will benefit the account holder.

In short, the financial industry, renowned for its ability to innovate and evolve, is likely to adapt to new regulation. Some providers may adjust their processes and recordkeeping to take advantage of carve-outs and exemptions. Perhaps more importantly, new technology and such products as target date funds are driving down the cost of serving small accounts and allow for modified account types or price structures. That trend is already well on its way.

Common Theme 3: Lack of Advice Will Cause Investors to Make Mistakes and Save Less

The third common theme is the argument that reduced professional advice will cause IRA investors to make more investment mistakes and save less. The argument tends to be based on research that purports to show benefits of financial advice. For example, ICI (2015a, p. 8): "Research shows that investors with access to advice have more diversified portfolios and take on more appropriate levels of risk than those who do not receive advice or information"; Litan and Singer (2015a, pp. 10-11): "brokers encourage their clients to save [... and] brokers help reduce investors' tendency to under-diversify in local stocks by overcoming the home-bias effect"; Quantria (2015, p. 12): "Access to financial advice counters the effects of a lack of financial literacy."

We agree that many investors benefit from professional advice, such as through increased saving or reduced investment mistakes. However, the benefits likely depend on the type of advice that is given. The studies under review are concerned that investors with small balances in brokerage accounts will lose access to advice. Such advice is subject to conflicts of interest and it is given to investors with relatively few assets. Under those conditions, much of the evidence put forth by the studies under review dissipates.

First, the studies tend to confuse the benefits of conflicted and non-conflicted advice. The studies that are cited as evidence of the benefits of professional advice tend to focus on non-conflicted advice. Each study cites its own body of literature and we will address specific citations in the individual reviews below, but our overall finding is that no study identified benefits of conflicted advice. (This does not imply that conflicted advice yields no benefits; we find only that none of the studies helped identify or quantify such benefits.)

Second, the studies tend to confuse causality with correlation. For example, Oliver Wyman (2015, p. 2) finds that "advised individuals aged 35-54 years making less than \$100K per year had 51% more assets than similar non-advised investors." It is doubtful that advisers deserve all the credit for the observed difference: did advisers prod their clients to save more, or are individuals who are serious about retirement saving more likely to seek professional advice? The causality may well run both ways. Oliver Wyman (2015) even designed and fielded its own survey but did not document the timing of financial advisers' involvement or any other questions that could have demarcated their role. Similarly, the literature cited in other studies under review did not distinguish causality from correlation.¹⁴

Third, the studies focus on gross benefits only, without taking costs into consideration. For example, some studies tout that advised households rebalance their portfolio more often than non-advised households. While that may seem laudable, rebalancing involves selling and buying securities and thus transaction

¹⁴ That said, two academic articles (both cited by NERA 2015a) presented evidence indicating that at least some of the increased saving occurred after an adviser was retained. Both related to overseas households and at least one to non-conflicted advice. Also see Section 3.

costs. Without information on brokerage commissions and front-end load fees, the net benefits of frequent rebalancing are questionable.¹⁵

Fourth, several studies reference DOL's (2011) analysis of expanded investment advice for evidence that advice results in investors' avoiding investment mistakes that would cost them roughly \$7bn-\$18bn. The focus of DOL (2011) is on *non-conflicted* advice. Also, the benefits estimated in that analysis relate to all IRA and DC balances. In contrast, the advice at issue here relates to only IRA balances that are too small to be served in an advisory account. While many IRA accounts are small, most of the dollars (and potential dollars of investment mistakes) are in larger accounts. Households with IRA assets under \$25,000 jointly own only 2% of total IRA assets (Panis and Brien, 2016).

In short, the studies under review cite numerous articles to demonstrate the benefits of advice, but none applies to the conflicted advice that is the focus of the Proposed Rule. We find it plausible that conflicted advice generates certain benefits, but their nature and magnitude remain in question.

Common Theme 4: The Academic Literature Is Misapplied

The fourth common theme is the argument that the RIA misapplies or misinterprets various academic studies upon which it relies. Examples include the following: "The academic research cited in the RIA is misapplied" (NERA 2015a, p. ii); "The academic studies the RIA cites do not support its sweeping claims" (ICI 2015a, p. 8); and "The RIA misuses these studies, however, and in the process, substantially overstates any benefits claimed from them" (Litan Singer 2015a, p. 22).

We carefully evaluated the commentators' arguments and found them to be lacking. This section discusses some recurring criticisms.

First, several studies argue that the academic studies are inapplicable because they are based on obsolete data. Christoffersen et al. (2013) used data from 1993 to 2009, Bergstresser et al. (2009) is based on data from 1996 to 2004, Del Guercio and Reuter (2014) covered 1992 through 2004, et cetera. The general argument is that load charges on mutual funds have diminished since the early 2000s, that conflicts of interest have faded correspondingly, and the underperformance of broker-sold funds found in the academic literature should no longer be of much concern.

The comments do not consider the fact that econometric relationships can be robust to changes in the levels of explanatory variables. For example, Christoffersen et al. (2013) estimated the relationship between broker payments and rates of return. Diminishing average loads and average broker payments do not affect the estimated relationship; a decline in broker payments would imply only that the underperformance became smaller. Indeed, ICI (2015c, p. 9) replicate Christoffersen et al.'s (2013) calculations with more recent data and find very similar (and even somewhat stronger) results. The average broker payment in the data of Christoffersen et al. (2013) was 2.3%, but the RIA adopted 1.41% for 2015 and assumed that it would continue to decline to 0.78% by 2036 (DOL 2015, p. 113).

¹⁵ Vanguard (2014, p. 15) made a similar observation.

Separately, the reduction in loads has not been so precipitous as some suggest. ICI (2015a, p. 9) argued that “in 2000 only about half of the funds with a front-end load share class also had no-load share classes [...]. By 2010, however, 90 percent of funds with a front-end load share class also offered a no-load share class.” However, as also pointed out by Christoffersen and Evans (2015), more important than a simple count of funds with no-load share classes are the dollars in those funds. Load funds accounted for 36% of retail assets in 2014, down from 49% in 2005 (ICI 2015d, Figure 5.10)—a smaller but far from negligible fraction, and still accounting for \$2.6 trillion dollars. Similarly, NERA (2015a, p. 31) claims that “[o]ver the period 1990-2013, front-end sales loads have declined by nearly 75 percent for equity funds and hybrid funds, and even more than that for bond funds.” As noted by ICI (2015d) and cited by NERA (2015a), part of this decline is due to load waivers for DC plan purchases, which are not at issue here. Indeed, DC assets rose nine-fold from 0.7 trillion in 1990 to 6.3 trillion in 2013 (ICI 2006, 2015d), thereby bringing down average load charges. In other words, the NERA claim exaggerates the decline in average load charges among mutual funds in IRAs.

Second, some studies objected to the fact that various academic analyses were not weighted by assets or by sales (except for Bergstresser et al., 2009). Indeed, for some purposes, weighting can be important. The average load charge across all load funds, for example, may be more meaningful when a weighted average is calculated. In econometric models, weights may affect the standard errors (the precision of the estimates). However, if the model is correctly specified, and especially if it controls for fund sales or assets, omitting weights does not introduce any bias in the parameter estimates.

Third, some studies interpreted the findings of Christoffersen et al. (2013) as evidence that funds with above-average broker payments underperformed direct-sold funds, and objected to an extrapolation to all funds with any broker payments. However, this objection rests on a logical error. Christoffersen and co-authors were correct to apply their results to all funds with broker payments, as was the DOL in its RIA. The authors estimated the relationship between broker payments (relative to an average) and rate of return. They found that performance decreases as broker payments increase; above-average broker payments suffer from above-average underperformance, and below-average broker payments suffer from below-average underperformance. Their measurement of broker payments relative to an average has no bearing on the estimated relationship for funds with below-average broker payments. Their estimates implied that funds with below-average broker payments had below-average underperformance, not that they had zero underperformance.

Another way to approach the argument is as follows. Christoffersen et al. (2013) found that funds’ rates of return decrease by 0.4972 percentage points for every percentage point increase in broker payments. Logically, funds with zero broker payments are free from a conflict of interest and their underperformance is zero. At a broker payment of 1%, estimated underperformance is $0.4972 \times 1\% = 0.50\%$, at 2% it is $0.4972 \times 2\% = 0.99\%$, etc. Christoffersen et al. (2013) reported average broker payments of 2.3%, where the estimated underperformance is $0.4972 \times 2.3\% = 1.14\%$ (reflects rounding error; the authors reported 1.13%). Funds with broker payments in excess of 2.3% underperform by more than 1.13% and funds with broker payments under 2.3% underperform by less than 1.13%; on average, load funds underperform by 1.13%.

Fourth, at least one study argues that the cited literature focuses on mutual funds, yet the DOL applies the results more widely. Indeed, the Proposed Rule applies to, for example, variable annuities that are purchased with IRA assets. Variable annuities offer sizeable sales commissions to brokers and thus present conflicts of interest. If anything, the conflicts of interest presented by annuities appear to be sharper than those of mutual funds. According to an industry expert cited in Scism (2012), commissions on indexed annuities average 6.3% of the principal payment, much higher than even the maximum front end load on most mutual funds.¹⁶

Fifth, some studies under review assert that the academic literature cited in the RIA is not applicable because it does not compare the costs and benefits of fiduciary accounts with those of brokerage accounts. This perspective is overly narrow and misses the point. The cited literature compares outcomes related to conflicted and non-conflicted compensation. Indeed, this distinction tends to correspond to the distinction between brokerage and fiduciary accounts. However, at issue is the conflicted compensation, not the name or structure of the account.

¹⁶ Also see AXA Equitable Life Insurance Company (2013), Jackson National Life Insurance Company (2013), and UBS (2015).

3. NERA ECONOMIC CONSULTING

Summary

In response to DOL's Proposed Rule, SIFMA submitted comments including "Comment on the Department of Labor Proposal and Regulatory Impact Analysis" by NERA Economic Consulting (NERA 2015a). In response to questions from the DOL, NERA provided additional details in a memorandum (NERA 2015b). This section contains a review of the NERA comment and the subsequent memorandum.

The basic premise of the NERA comment is that the Proposed Rule will force brokerage accounts to close, with two consequences: accounts with sufficiently high balances will be moved to a fee-based model, and account holders with small balances will lose access to advice. For the first group, NERA addresses costs of impeding the commission-based investment model. For the second group, it addresses costs of losing access to advice. Finally, NERA challenges the DOL's estimates of the costs of conflicted investment advice.

NERA's analysis of the costs of impeding the commission-based investment model is based on a confidential data set of IRAs which is not publicly available. Hence it is not possible to critically assess some important dimensions of NERA's calculations.

NERA argues that commission-based accounts incur lower fees than fee-based accounts. However, its comparison excludes important fee components. NERA acknowledged the bias but did not do anything to mitigate it even though that would have been relatively straightforward with publicly available information. Separately, and contrary to NERA's claims, NERA's own data suggest that commission-based accounts may have underperformed fee-based accounts on a risk-adjusted basis. In particular, rather than earning virtually the same median return as fee-based accounts, because brokerage accounts in NERA's database may have been riskier, they could have earned *higher* returns, reflecting the risk premium that should have been earned by riskier assets during the period of a strongly appreciating overall market between mid-2012 and early 2015. A particularly troublesome aspect of NERA's analyses is the lack of detail about its data source. NERA (2015a) only presents findings for median accounts. Concerns over conflicted advice are likely to manifest themselves away from the median: commissions may be excessive for a minority of accounts, excessive trading may be found in a minority of accounts, underperformance may be serious for a minority of accounts, etc. Median statistics cannot show any such pattern. In addition, even after being asked directly and given the opportunity to investigate, NERA (2015b) could not provide assurances that the data were representative of financial institutions or IRA accounts in the United States.

NERA next sets out to show that loss of professional advice would be detrimental to investors. We find it plausible that many advisers help reduce investment errors, but the evidence put forth by NERA is not convincing—the mostly foreign studies reviewed may not be applicable to the U.S. context, the studies are selectively quoted or even misquoted, NERA highlights only benefits of advice without weighing those against their costs, and some studies confuse correlation with causality. That said, two studies provided compelling evidence of value added by advisers. One of those articles related to advisers who were relatively free from conflicts of interest,

confirming the value of fiduciary advice but not helping the case for conflicted advice. Separately, NERA misquotes DOL's own analysis of losses due to investment errors. DOL's figures applied to fiduciary advice for all DC plan participants and IRA holders, whereas NERA is concerned with non-fiduciary advice for IRAs with small balances—NERA's assumption that the value of non-fiduciary advice is the same as that of fiduciary advice ignores the very impetus of the Proposed Rule, and even if the two types of advice were equally effective at avoiding investment mistakes, NERA should have reduced the purported benefits by at least 98% to account for the much smaller asset base.

Finally, the NERA comment challenges DOL's calculations of the cost of conflicted advice, asserting that the DOL misapplied or misinterpreted academic studies. Each of its lines of attack however, falls flat due to NERA's own misunderstanding of the literature and of DOL's approach and due to NERA's deficient and narrow interpretation of the applicability of academic studies.

Synopsis

The basic premise of the NERA comment is that the Proposed Rule will force brokerage accounts to close, with two consequences: accounts with sufficiently high balances will be moved to a fee-based model, and account holders with small balances will lose access to advice. For the first group, NERA addresses costs of impeding the commission-based investment model (Section 1). For the second group, it addresses costs of losing access to advice (Section 2). Finally, NERA challenges the DOL's estimates of the costs of conflicted investment advice (Section 3).

To analyze costs of impeding the commission-based investment model, NERA collected account-level data on over 63,000 fee-based and commission-based IRAs. The authors found that median fees on fee-based accounts were 57-101 bps higher than on commission-based accounts, depending on account balance. They also found that, at the median, fee-based accounts traded more frequently than commission-based accounts. Based on median quarterly rates of return, the authors found that commission-based accounts did not underperform fee-based accounts between mid-2012 and early 2015.

About 40% of commission-based IRA balances were below \$25,000, characterized as the "conservative minimum account balance" (NERA 2015a, p. 9) required to open a fee-based account. NERA assumes that the Proposed Rule will trigger a loss of access to financial advice for these account holders. NERA asserts that loss of professional advice would cost more than the current cost of conflicted advice because individual investors benefit from financial advisers through better diversified portfolios, fewer investment mistakes, tax minimization, increased savings, and lower cost of information. NERA then reviews a 2011 analysis by the SEC into potential consequences of imposing fiduciary duty on brokers and a 2011 analysis by the DOL, which estimated that more advice to DC plan participants and IRA holders could prevent mistakes that would cost investors \$7 billion to \$18 billion annually.

The third section concerns the cost of conflicted investment advice. Referring to estimates of the cost of conflicted advice in the Proposed Rule's RIA, NERA concludes that the "range of numbers is so wide as to provide no scientific confidence in the DOL's own methodology" (NERA 2015a, p. 30). It then argues that the RIA misapplied findings of the academic literature on the cost of conflicted advice: the

literature focuses on mutual funds, but the RIA applies it also to variable annuities and other products; the RIA takes results associated with higher-than-average load funds and misapplies them to all funds; and the literature does not compare the costs and benefits of fiduciary accounts with those of brokerage accounts.

Discussion

Overview

Throughout its comment, NERA discusses and compares fee-based account and commission-based accounts. "Fee-based accounts are charged a fixed fee as a percentage of assets whereas commission-based accounts are charged fees based on trading and other activity" (NERA 2015a, p. 3). In related literature, fee-based and commission-based accounts are also referred to as advisory and brokerage accounts, respectively. Advisers to fee-based advisory accounts are generally held to a fiduciary standard of conduct, whereas advisers to commission-based brokerage accounts are held to a lower suitability standard. The Proposed Rule is concerned that advisers to commission-based accounts face conflicts of interest.

The remainder of this section follows the organization of the NERA comment. We first discuss NERA's Section I on the costs of impeding the commission-based investment model, then Section II on costs of losing access to advice, and finally Section III on the costs of conflicted advice.

I. Costs of Impeding the Commission-Based Investment Model

As summarized above, this section compares commission-based and fee-based IRA accounts with respect to fees and rates of return. NERA's analysis is primarily based on a confidential dataset of over 63,000 IRA accounts with data ranging from 2012 through the first quarter of 2015.

The NERA comment purports to show that fee-based accounts are more expensive than commission-based accounts. The magnitude of the difference ranges from "about 57 basis points (bps) for relatively small accounts (those with balances below \$25,000) up to about 1 percent for accounts with balances from \$100,000 to \$250,000" (NERA 2015a, p. 6). But fees in this comparison "exclude revenue that the firm may receive indirectly from the account-holder, such as markup/markdown revenue or 12b-1 fees" (NERA 2015a, p. 4). These and other indirect revenue components vary across products, tend to constitute conflicted compensation, and their exclusion therefore makes brokerage accounts appear less expensive than they really are. In the absence of these fees, it cannot be determined whether fee-based accounts are more or less expensive than commission-based accounts. For example, 12b-1 fees and shareholder service fees can run as high as 100 bps (SEC 2015); at that level, fee-based accounts would incur lower fees than commission-based accounts.

The NERA comment recognizes this deficiency in a footnote and seeks to address it in its comparison of rates of return, but not in its fee comparison.

The NERA memorandum defended the exclusion of indirect fees with the assertion that its data set did not contain information related to such fees. However, the detailed, account-level data that NERA compiled presumably included information on portfolio compositions, and 12b-1 fees for individual funds are readily available from

Morningstar and other sources. In other words, NERA's fee comparison is biased; NERA acknowledged the bias and did not do anything to mitigate it even though doing so would have been relatively straightforward with publicly available information.

NERA's Section 1 continues with the argument that individuals self-select into the account type that favors their behavior, based on more frequent trades in fee-based accounts than in commission-based accounts. But the comment fails to qualify what kind of trade transactions have been included and excluded from this comparison. For example, it is not clear whether fee payments for account maintenance and advisory services are included. For fee-based accounts, these fees are expected to be small and periodic and could skew the results. By contrast, commission-based accounts which have few direct fees assessed may not have such trades. Also, many trades may be related to purchases, (mandatory) distributions and dividend reinvestments.¹⁷ The NERA comment does not specify either whether these trades are removed from this analysis.

More generally, NERA provided very little explanation of its data source, which raises questions about the completeness and robustness of its findings. The analysis purports to compare commission-based and fee-based accounts, but the age distribution (Exhibit 1) and the account balance distribution (Table 1) are reported across all account holders. Further, the comparisons of account fees and trading frequency are carried out entirely on the basis of median values, which may not reflect relevant information on 63,000 IRAs. For example, NERA notes that "it is worth noting that the data does not seem to show 'churning,' the needless buying and selling of securities. We see the median commission-based account had traded 6 times in 2014. Such trading is more consistent with a buy-and-hold strategy than churning" (NERA 2015a, p. 8). We agree that the median number of trades does not reflect churning. However, the presented figures are also consistent with abundant churning among 49% of commission-based accounts. Insofar we are aware, nobody is alleging that conflicts of interest cause advisers to churn almost one-half of commission-based accounts, but the DOL would presumably be concerned if it occurred in 5% of the accounts. To that end, the 95th percentile of number of trades would be informative. Based solely on the median, NERA's conclusion that churning is not an issue is unconvincing.

NERA presented even fewer relevant details in its comparisons of rates of return for fee-based and commission-based accounts. This part of the analysis is highly relevant, because much of the concern over conflicts of interest is driven by underperformance of funds sold in commission-based accounts. This underperformance has been documented based on publicly available data in peer-reviewed academic articles (e.g., Bergstresser et al., 2009; Christoffersen et al., 2013; Del Guercio and Reuter, 2014). NERA claims to have found that rates of return in commission-based accounts are in fact about equal to those in fee-based accounts. For this claim—based on confidential data and without peer review—to be credible, the analysis needs to be extensively documented and stress-tested. Instead, NERA devotes merely one page to the analysis description and presents

¹⁷ The NERA memorandum showed that individuals age 60 and older are more prevalent among fee-based account holders than among commission-based account holders. These individuals may take regular distributions to fund their retirement, and may even be forced to take distributions because of minimum distribution requirements that apply above age 70½.

quarterly differences in median returns only (NERA 2015a, Table 4, p. 10), with no controls for such factors as the riskiness of investments that are prominent in the academic literature. Even the quarterly returns remain unreported; only the differences in median returns between fee-based and commission-based accounts are provided.

In response to questions from the DOL, the NERA memorandum provides some more, though still inadequate, details.

The analysis of rates of return falls short in several aspects.

First, the comment fails to adjust for differences in riskiness (volatility) of account portfolios. This is important if assets in fee-based and commission-based accounts differ in the average level of risks. For example, a portfolio invested only in stocks that make up the S&P 500 index would have realized compound annual growth rate of approximately 19% over the period of the study, much higher than the historical average rate of return on stocks (finance.yahoo.com, SP500TR). But of course investing in stocks only will not be suitable for all investors, particularly not for those nearing retirement. The NERA memorandum (NERA 2015b, p. 4) shows that account holders of fee-based accounts tend to be older than commission-based account holders. Roughly 58% of fee-based account holders are age 60 or older, compared with roughly 48% of commission-based account holders. Based on their higher ages, fee-based account holders probably invest in less risky assets than commission-based account holders.

The NERA memorandum suggests that commission-based accounts are invested in riskier assets than fee-based accounts. Table 1 below transcribes the 25th and 75th percentiles of quarterly rates of return for fee-based and commission-based accounts in the NERA sample, as provided in the NERA memorandum (NERA 2015b, p. 3). We calculated the interquartile range—the difference between the 75th and 25th percentiles—for each quarter. The interquartile range is a measure of the dispersion of rates of return, which may be related to the riskiness of invested assets in the individual accounts. In 10 of the 11 quarters of data, the interquartile range for commission-based accounts exceeded that of assets in fee-based accounts.

Table 1. Percentiles and Interquartile Range of Quarterly Rates of Return of Fee-Based and Commission-Based Accounts in the NERA Sample

Quarter	Fee-based accounts			Commission-based accounts		
	25th percentile	75th percentile	Interquartile range	25th percentil	75th percentil	Interquartile range
Jun-12-Sep-12	3.16%	5.45%	2.29%	2.58%	5.76%	3.18%
Sep-12-Dec-12	-1.16%	1.79%	2.95%	-0.93%	1.58%	2.51%
Dec-12-Mar-13	3.27%	7.81%	4.54%	3.44%	9.71%	6.27%
Mar-13-Jun-13	-1.76%	0.95%	2.71%	-0.90%	2.27%	3.17%
Jun-13-Sep-13	3.29%	6.44%	3.15%	1.45%	6.41%	4.96%
Sep-13-Dec-13	3.81%	7.14%	3.33%	2.51%	8.24%	5.73%
Dec-13-Mar-14	0.41%	1.77%	1.36%	0.26%	2.55%	2.29%
Mar-14-Jun-14	2.58%	4.17%	1.59%	2.01%	4.66%	2.65%
Jun-14-Sep-14	-2.52%	-0.80%	1.72%	-1.85%	0.18%	2.03%
Sep-14-Dec-14	0.19%	2.54%	2.35%	-0.11%	3.17%	3.28%
Dec-14-Mar-15	0.74%	2.53%	1.79%	0.00%	2.69%	2.69%

Source: NERA (2015b), p. 3.

In a bull market (such as the period studied by NERA), the rates of return on riskier assets may be expected to be higher, suggesting that commission-based accounts should have returned a premium over less risky assets in fee-based accounts. NERA (2015a) did not control for volatility or find such a premium. The limited information that NERA made available does not permit quantifying the risk premium that commission-based accounts should have earned, but it may explain why NERA did not find underperformance. Without exploring the issue, NERA (2015a, p. 11) had insufficient basis to conclude that “there is no support in this data for the contention that commission-based accounts underperform.”

Second, the NERA comment based its underperformance analysis entirely on median quarterly rates of return. At best, such data support a conclusion about underperformance at the median; they do not support any conclusion about accounts above or below the median. For example, the median would be the same if 49% of commission-based accounts performed extremely poorly. Again, insofar we are aware, nobody is alleging that conflicts of interest cause advisers to place almost one-half of commission-based accounts in grossly underperforming funds, but the DOL would presumably be concerned if it occurred in 5% of the accounts. To that end, the 5th percentile of rates of return would be informative. Based solely on the median, NERA’s conclusion that underperformance is not an issue is unconvincing.

The NERA sample raises many more questions. For example, some IRAs presumably included variable annuities; how were those treated in the analysis? Separately, there is no discussion of sampling weights or of any attempt to ensure representativeness of the sample. NERA’s response to DOL questions about representativeness of the sample and generalizability of the analysis findings was only that the “the sample accounts contained a wide variety of balances, transaction activity levels, and customer ages” (NERA 2015b, p. 1) and that the authors had “confidence that our data included a diverse selection of accounts, with no evidence of any bias in the data” (NERA, 2015b, p. 2). Of course it is impossible to detect a bias without conducting a comparative analysis of the IRA population.¹⁸ The validity of using a particular sample to reflect the characteristics of a population can be determined by describing the sampling process exactly and in some cases by assessing the characteristics of the sample compared to the population of interest. NERA provided none of this standard information. NERA did not even describe the firms at which the accounts were held, other than that they are SIFMA members.

Finally, NERA’s Section 1 was motivated by the assertion—without evidence or even arguments—that the Proposed Rule and associated RIA “have led many to conclude that the proposal would effectively make the commission-based brokerage model

¹⁸ For example, NERA (2015a, p. 5) reported an average IRA balance in its sample of \$174,034. In contrast, Panis and Brien (2016, p. 39) documented an average balance of \$100,998 in the nationally representative Survey of Consumer Finances (assets of \$6.676 trillion divided by 66.1 million accounts). Similarly, Copeland (2015, p. 7) documented an average IRA account balance of \$95,363 for 2013. The latter sources include small accounts, whereas NERA (2015a) reportedly excluded accounts under \$1,000. However, for the NERA average to be consistent with the national average, at least 42% of accounts must have been under \$1,000. According to ICI, only 22% of traditional IRAs (ICI 2015e) and 24% of Roth IRAs (ICI 2015f) had balances under \$5,000 in 2013. In other words, NERA’s sample is biased toward larger accounts.

unworkable for investment accounts covered by ERISA and similar sections of the IRS code” (NERA 2015a, p. 2). NERA proceeds to assert that commission-based “investors will have to move to fee-based accounts or lose access to professional investment advice entirely” (NERA 2015a, p. 9). As discussed in Section 2 (Common Themes), the comment ignores the possibility that financial institutions will modify their commission-based account types or introduce types other than current-style commission-based or fee-based accounts.

II. Cost of Losing Access to Advice

The second section of the NERA comment focuses on the cost of losing advice. This issue may become relevant if future regulation prompts financial institutions to discontinue certain IRAs. NERA asserts—without evidence—that IRAs with a balance of less than \$25,000 may no longer receive advice. Based on its proprietary dataset of IRA accounts, NERA (2015a) projects that 40% of commission-based accounts will lose access to a financial adviser. NERA’s database is reportedly drawn from SIFMA members, which include discount brokerages. To the extent the IRA accounts include discount brokerage accounts, the fraction of accounts that will lose access to advice is in fact lower.

The assertion that the commission-based brokerage model will become unworkable is questionable and has been addressed in the Section 2 of this report (Common Theme 2).

We first narrow down the area of interest. The Proposed Rule aims to mitigate conflicted advice. For the purpose of quantifying the cost of losing access to advice due to the Proposed Rule, the focus should therefore be on benefits of conflicted advice only. It is not relevant or logical to discuss the benefits of fiduciary advice, because it will not be reduced or restricted by the Proposed Rule.

NERA’s comment references a 2011 DOL analysis of losses due to investment errors to suggest that DOL itself attributed huge value to professional advice.¹⁹ NERA states that “the DOL estimated that participant-directed retirement savings account holders make investment mistakes *in the absence of professional advice* valued at an aggregate of “more than \$114 billion in 2010” (NERA 2015a, pp. 11-12; emphasis added). However, NERA misinterprets and misapplies DOL’s earlier analysis. The wording suggests that professional advice could prevent \$114 billion in losses, but the DOL in fact estimated that increased access to advice would reduce these losses by \$7 billion to \$18 billion. Further, those estimates related to assets in all DC plans and IRAs, rather than just IRAs with balances under \$25,000, which account for less than 2% of total DC plan and IRA assets (Panis and Brien 2016). Finally, the estimated reduction of investment errors would be the result of increased access to fiduciary advice, not conflicted advice.

NERA cites a number of studies that found that many individual investors make suboptimal investment decisions: they may be more inclined to lock in gains than to cut losses, and they may trade too often and incur excess transaction costs. The authors then discuss a number of articles that they claim demonstrate that financial advisers help reduce investment errors. While we find it plausible that many advisers help reduce investment errors, the evidence put forth by NERA is not convincing—the mostly foreign studies may not be applicable to the U.S. context, the studies are

¹⁹ See Federal Register, Volume 76, pages 66136-66167 for the 2011 DOL analysis.

selectively quoted or even misquoted, NERA highlights only benefits of advice without weighing those against their costs, and some studies confuse correlation with causality.

First, consider applicability of the cited studies. The NERA comment's section on "Benefits of Financial Advisors" (NERA 2015a, pp. 17-22) discusses 17 papers. Eleven papers are based on foreign data (Germany, Canada, Netherlands, United Kingdom, Israel, and Australia), two were theoretical exercises without empirical data, and only four were based on U.S. data. Legal and regulatory regimes vary by country; therefore, the foreign studies are relevant only to the extent that foreign advisers are subject to conflicts of interest. For example, Bluethgen, Gintschel, Hackethal and Mueller (2008) state that "If retail financial advisory services differ across countries in terms of cost and quality then their effects on household portfolios might also be very different" and go on to suggest that "Regulation aiming to enhance investor protection should then not only focus on capital markets themselves but also set and enforce minimum quality standards for financial advisory services," something the DOL proposal seeks to achieve. However, NERA offers no discussion of foreign advisers' conflicts of interest, if any, and it offers no discussion of the standards—fiduciary, suitable, or otherwise—to which foreign advisers are held. Even for the U.S.-based studies, it is unclear whether the advice under analysis was conflicted.

Second, NERA selectively quotes from the papers it reviews and omits essential findings that counterbalance claims of adviser value. For example, NERA (2015a, p. 17) quotes Gerhardt and Hackethal (2009) as finding "that there are clearly positive effects to working with an advisor." However, the primary conclusion of Gerhardt and Hackethal (2009, p. 22) was that "major aspects of the (positive) effects that have been attributed to the influence of professional investment advisors is in fact due to differences in investors' behavior. The actual effect of investment advisors is – while clearly existing – much smaller than assumed by previous studies." In another example, based on Montmarquette and Viennot-Briot (2012), The Investment Funds Institute of Canada (2012) is quoted as noting "that research proves that advice has a positive and significant impact on wealth accumulation" (NERA 2015a, p. 18). However, the paper does not compare rates of return earned by advised and non-advised investors. The only "highly plausible explanation" offered for observed differences in wealth is that advised households save more than non-advised households, and advisers may not deserve full credit for that relationship (see below).²⁰ In a third example, Kramer (2012) is summarized as finding that "advised portfolios are more diversified and perform better than self-directed portfolios" (NERA 2015a, p. 18). However, Kramer (2012, p. 395) in fact found "no evidence of differences in risk-adjusted performance." In a fourth example, "Kinniry, Jaconetti, DiJoseph and Zilbering (2014), argue that [...] advisors can potentially add about 3 percent in net returns to investors" (NERA 2015a, p. 18). The key term here is "potentially": this paper, which NERA characterizes as "widely-cited," is marketing and training material for advisory services of a large financial service provider. It describes "best practices" (which almost by definition not all conflicted advisors employ) and offers primarily examples, rather than empirical evidence for its estimates of behavioral biases.

²⁰ NERA does cite that finding on page 20, but incorrectly added that Montmarquette and Viennot-Briot also pointed at improved asset selection as a highly plausible explanation. In fact, Montmarquette and Viennot-Briot attributed only increased savings to the presence of advice.

Third, with worthy exceptions noted below, most of the studies reviewed by NERA fail to weigh the benefits of advice against their costs or gloss over causality issues. For example, several studies found that advised portfolios were better diversified than non-advised ones. We agree that, all else equal, a well-diversified portfolio is generally preferable over a highly concentrated portfolio. However, the advice and the diversification are not free of charge, and a central issue is whether the diversification as advised by financial advisers generated long-term benefits. Despite NERA's repeated claims that advisers help investors make better investment decisions, there is no evidence that advised portfolios outperform non-advised portfolios. Separately, several studies showed that advised households save more than non-advised households. However, it is doubtful that advisers deserve all the credit for observed differences: did advisers prod their clients to save more, or are individuals who are serious about retirement saving more likely to seek professional advice? The causality may well run both ways.

There are two noteworthy exceptions to the causality defects of many studies that purport to measure the value of advisers. The first is Montmarquette and Viennot-Briot (2012), also published as Montmarquette and Viennot-Briot (2015). The authors used Canadian data with information on whether respondents used a financial adviser and the "tenure" of advice, i.e., how long they had been consulting an adviser. The information on adviser tenure reduces some causality issues. For example, while many studies have documented that advised households tend to be wealthier than non-advised households, the authors found that the wealth disparity increased with the duration over which they had been advised. It is still possible that people who are serious about preparing for retirement are more likely to consult an adviser, but Montmarquette and Viennot-Briot (2015) provide convincing evidence that advisers add value by helping people save more.

The other exception is Kramer (2012), who used Dutch data with information both before and after investors started consulting an adviser. He found that portfolios became better diversified after "advisory intervention." Despite better investment behavior, he did not find statistically significant differences in rates of return between advised and non-advised accounts. The financial advisers in his "sample are paid fixed wages only, so they have no direct personal financial incentive to generate commissions, but career and prestige considerations are likely to play a role," suggesting that the value they added was in an environment that was relatively free of conflicts of interest. He also noted that the 5th percentile of the advised portfolio value distribution was €600, suggesting that relatively conflict-free advice can be available even at low account balances.

III. The Cost of Conflicted Investment Advice

The last section of the NERA comment focuses on estimates of harm caused by conflicted advice put forth in the DOL's RIA.

The NERA authors take issue with the fact that the RIA presents many different estimates. However, they fail to recognize that the RIA, given the uncertainty embedded within many of the assumptions, has adopted a scenario-based analysis to present both conservative and likely estimates of the harm caused by conflicted advice. This is considered a best practice when uncertainty in assumptions is involved and is widely used when forecasting into the future (e.g., International Actuarial Association, 2013; Maack 2001). Also in light of its sensitivity analyses and

its extensive discussion of uncertainty, the 243-page RIA demonstrates a thoroughness that is commendable.

The NERA comment also asserts that the RIA misapplies the academic literature. These assertions have been addressed in Section 2. In particular, the NERA review finds that “The academic literature cited in the RIA does not compare the costs and benefits of fiduciary accounts with those of brokerage accounts” (NERA 2015a, p. 33). Indeed, such data have not been available, at least not to date and not publicly. NERA appears to have access to IRA data that permit a comparison of fiduciary and brokerage accounts. Our reading of very crude summary statistics of those data indicate that brokerage accounts likely underperformed fiduciary accounts on a risk-adjusted basis—see the discussion above. That aside, the academic literature has centered on underperformance due to conflicts of interest, which is precisely the target of DOL’s Conflict of Interest Proposed Rule. Finally, NERA incorrectly states that DOL has misapplied Christoffersen et al.’s results:

In particular, their study finds evidence that a subset of funds, those whose front-end loads are higher than other funds with similar characteristics, underperformed the average return of their fund category during the next year. In formulating much of their “cost of conflicted advice” aggregate figures, the DOL then assumes that *all* IRAs invested in front-end load funds suffer the same underperformance, thereby mistakenly applying a result from a subset of load funds to all load funds.

The extrapolation the DOL made is analogous to the following: Suppose we conduct medical research and find that people who consume more salt than average have a lower life expectancy by five years, and we then conclude that eating no salt will increase the life expectancy of everyone by five years. This is a logical fallacy. We have no evidence that people who eat a “normal” amount of salt would benefit from reduced salt intake, and so extrapolating to them is an error in logic. (NERA 2015a, pp. 32-33.)

NERA’s analogy does not describe how DOL has applied Christoffersen et al.’s relationship, which indicates that a reduction in front-end loads increases returns, regardless of whether the load is above or below average. In the terms of NERA’s analogy, people who eat a “normal” (or even less than normal) amount of salt would in fact benefit from reduced salt intake.

4. OLIVER WYMAN

Summary

In response to DOL's Proposed Rule, several financial firms submitted a report titled "The role of financial advisors in the U.S. retirement market" (Oliver Wyman, 2015). This section contains a review of the Oliver Wyman study.

The Oliver Wyman study establishes that financial advisers tend to be involved when events occur or circumstances exist that are good for retirement security: small businesses sponsor employee retirement plans, individual investors are wealthier, individual investors' portfolios are more diversified, et cetera. In a major shortcoming, the study credits financial advisers for progress toward retirement security in which they were not involved. Attributing all observed differences between advised and non-advised businesses or individuals to financial advisers is clearly an overstatement; small businesses may have retained an adviser after deciding to set up a retirement plan, investors may have retained an adviser after accumulating substantial assets, et cetera. Quite plausibly, causality runs in both directions: some advisers foster retirement security, and some advisers get involved with businesses or individuals who have already made progress toward retirement security. The latter advisers may still add value, but did not play a role in what occurred before their involvement.

Oliver Wyman designed its own surveys of small businesses and individual investors, but missed an opportunity to document the contributions of financial advisers as opposed to progress toward retirement security without the involvement of advisers. The study does not report on the timing of financial advisers' involvement or any other questions that could have demarcated their role. Further, the study does not distinguish between broker-provided and fee-based advice, instead treating them the same and failing to acknowledge that the Proposed Rule targets conflicted advice only.

In another major shortcoming, the study does not address the costs of financial advice to small businesses or individual investors. For example, the study shows that advised individuals rebalance their portfolios more often than non-advised individuals. While that may seem laudable, rebalancing involves selling and buying securities and thus transaction costs. Without information on brokerage commissions and front-end load fees, the *net* benefits of frequent rebalancing—precisely the type of issue that the Proposed Rule aims to address—cannot be determined.

Based on unspecified other sources, the Oliver Wyman study contends that the Proposed Rule will likely limit the ability of financial advisers to offer services to small businesses and individual investors, and raise the cost of such services. The study fails to establish to what extent financial advisers deserve credit for favorable outcomes, fails to examine whether costs outweigh purported benefits of financial advice, and fails to consider whether alternative advisory mechanisms could emerge to serve affected investors.

Synopsis

The Oliver Wyman study evaluates the role of financial advisers in two areas: advising companies on how to set up DC plans and advising individuals on retirement saving. The first part is based on a survey, conducted by Oliver Wyman, of about 1,200 small businesses. The second part is based on another survey, also conducted by Oliver Wyman, of about 4,400 retail investors and also on data from a third party, IXI Services, on consumer investments. We understand those investments data to be aggregated, i.e., without account-level details.

The first part of the study found that financial advisers assist business owners with setting up a DC plan for their employees. "Specifically, businesses with 1–9 employees with a financial advisor are almost twice as likely to set up a retirement plan as are businesses without financial advisors (51% vs. 26%). Businesses with 10–49 employees with a financial advisor are 48% more likely (77% vs. 52%) and businesses between 50 and 100 employees are 19% more likely (89% vs. 75%) to set up a plan" (Oliver Wyman 2015, p. 14). The study asserts that the Proposed Rule would force financial advisers to stop providing retirement plan services to small businesses; "many small businesses are likely to close or not open plans due to the additional administrative burden as a result" (Oliver Wyman 2015, p. 38).

The second part of the Oliver Wyman study found that investors with a financial adviser had more financial assets and exhibited better investment behavior along several dimensions than investors without a financial adviser. The study asserts that the Proposed Rule would reduce access to financial advisers by retail investors, who would face higher expenses to maintain access to advisers or, deprived of advice, would save less for retirement and would invest less wisely.

In summary, the Oliver Wyman study concludes that the DOL's Proposed Rule would likely reduce retirement savings.

Discussion

Overview

The Oliver Wyman study attempts to relate the message that financial advisers make good things happen: small businesses set up DC plans, and individual investors accumulate retirement assets and invest wisely. But the study's approach and its findings do not support that message. The study demonstrates a correlation between the involvement of advisers and favorable retirement security outcomes, but it does not demonstrate a causal link. Did small businesses set up DC plans because financial advisers convinced them to do so, or did small businesses decide to set up a DC plan before retaining a financial adviser? Did investors accumulate substantial wealth because financial advisers prodded them to save more and invest wisely, or did investors decide to seek professional advice after accumulating substantial wealth? In addition, the study does not discuss how advisers are compensated, i.e., does not identify the contributions of conflicted advice. The study is based on Oliver Wyman's own surveys of small businesses and retail investors. Either Oliver Wyman missed an opportunity to directly ask about compensation, causality, and timing, or the surveys—which to our knowledge are not publicly released—yielded relevant insights which the study's authors did not describe.

In its discussion of the report's implications, the authors assert that the Proposed Rule would likely reduce access to financial advisers by small businesses and retirement investors.²¹ The authors argue that as a result, "The benefits financial advisers provide are now at risk" (Oliver Wyman 2015, p. 37). However, since the study did not establish to what extent financial adviser involvement *causes* favorable outcomes or to what extent the purported benefits related to conflicted advice, it is not clear what those benefits are.

A second issue relates to costs. Suppose the Proposed Rule would limit access to financial advisers, as the Oliver Wyman study asserts. While the Oliver Wyman study did not distinguish correlation from causality, it seems entirely plausible that the services of financial advisers bring certain benefits. Directly or indirectly, today's advisee pays for the services of financial advisers. Fewer services thus imply both a loss of benefits and a reduction of costs. The Oliver Wyman study highlights the loss of benefits, but does not discuss costs. In a complete evaluation, the loss of benefits would be weighed against lower costs for advisees.

A third issue concerns the static view of the world that the Oliver Wyman study adopts. Suppose, as asserted, that financial advisers would cease to advise certain clients. In the Oliver Wyman view, these clients would face undesirable consequences: "many small businesses are likely to close or not open plans" (Oliver Wyman 2015, p. 38), "Individuals are less likely to open an IRA, leading to lower savings rates and increased cash-outs when changing jobs" (Oliver Wyman 2015, p. 39), and "Unadvised individuals are likely to carry excess portfolio risk due to less diversification and less frequent re-balancing compared with advised individuals" (Oliver Wyman 2015, p. 39). The study ignores the possibility that the abandoned clients would find another financial adviser. For example, in a preliminary evaluation of changes to the financial system in the United Kingdom, which recently banned payments to financial advisers that depend on the advice given, Europe Economics (2014, p. 63) found:

Some advisers have sought to terminate unprofitable client relationships. Data from NMG Consulting, for example, imply that in the year to Q1 2014 about 310,000 clients stopped being served for this reason. On the other hand 820,000 clients were gained in the same period. The same survey indicates that advisers refused to serve about 60,000 (potential new) clients in the same period. If we assume that many of those clients with relationships terminated on the grounds of inadequate profitability sought out another adviser, the positive net increase in customers served suggests that such

²¹ The study itself does not provide evidence that access to financial advisers would be reduced. Instead, the study relies on unspecified other sources. For example (Oliver Wyman 2015, pp. 2-3):

"Many stakeholders are now analyzing the technical details of the newly proposed rule, and there is growing concern that the proposal would again result in unintended consequences, including limiting the ability of financial services firms and individual financial advisors to offer services to individual IRA holders and small businesses, as well as increasing investor costs due to new expenses associated with implementing the rule and transitioning many clients to a higher cost advisory model."

looking around for a replacement was largely successful. We cannot rule out the existence of a residual group of consumers denied service in this way. However these data do not speak to a significant issue here.

In other words, the industry adjusted to the payments ban and some adviser-client relationships were reshuffled. The initial U.K. experiences suggest that the Proposed Rule would not deprive many retail investors of financial advice.

Similarly, the Oliver Wyman study argues that the Proposed Rule would make it difficult for financial services firms to offer brokerage accounts and would migrate accounts to more expensive advisory accounts. "Almost all retail investors would face increased costs (73% to 196% on average) from providers shifting clients to a fee-based advisory model"²² (Oliver Wyman 2015, pp. 7 and 38). Again, this view is overly static, apparently grounded in the premise that all accounts, after regulatory changes, will continue to be either traditional brokerage accounts or traditional advisory accounts, with traditional fee structures and traditional product offerings. The fee comparison fails to capture all differences between brokerage accounts and advisory accounts, as demonstrated by the fact that some investors opt for an advisory account today, with its reportedly higher fees. The financial services industry may well adapt and make changes to brokerage accounts or develop an alternative account type. The account fees will presumably be more transparent than they are in today's brokerage accounts and some clients may be unpleasantly surprised at the expense, but they will be no worse off than in their current brokerage account. If anything, retirement investors will be in a better position to manage their expenses.

Please refer to Section 2 for additional discussion of common themes that apply to the Oliver Wyman study.

The remainder of this section discusses Parts I and II of the Oliver Wyman study, related to small businesses and retail investors, respectively, followed by a discussion of issues with the data sources of the study.

I. Role of Financial Advisors in the Defined Contribution Plan Market

Part I of the Oliver Wyman study starts with statistics on the large and increasing role of DC plans for retirement financing and demonstrates that smaller firms are less likely to sponsor a pension plan than larger firms. These patterns are widely known and not controversial.

The study continues with results from the Oliver Wyman Small Business Survey 2014, a survey of owners and human resources (HR) decision makers at payroll-

²² The increased costs figures are based on Oliver Wyman (2011). They reportedly exclude "marketing and distribution, shareholder services, and other fees not directly paid by investors" (Oliver Wyman 2011, p. 22). The exclusion of marketing and distribution fees is puzzling and skews the results. For example, front-end load fees are disproportionately present in brokerage accounts. Indeed, Oliver Wyman's increased cost figures seem counter intuitive as they suggest that the largest accounts would incur the greatest cost increases—both in dollars and percentage terms.

based businesses with between 1 and 100 employees. As noted in the study, the survey had a sample size of 1,216 valid complete responses.

A key finding of the Oliver Wyman study comes from a comparison of retirement plan sponsorship among small businesses that did or did not consult with a financial adviser: "We found that 41% of small businesses with 100 or fewer employees work with a financial advisor, and that these firms are significantly more likely to set up a retirement plan. Specifically, businesses with 1–9 employees with a financial advisor are almost twice as likely to set up a retirement plan as are businesses without financial advisors (51% vs. 26%). Businesses with 10–49 employees with a financial advisor are 48% more likely (77% vs. 52%) and businesses between 50 and 100 employees are 19% more likely (89% vs. 75%) to set up a plan" (Oliver Wyman 2015, p. 14).

Oliver Wyman's key finding may be misleading for several reasons. First, the study speaks of advised small businesses being more likely to "set up" a retirement plan and shows "plan formation rates" (Oliver Wyman 2015, p. 14). This phrasing suggests starting or initializing a plan once an adviser becomes involved. However, the survey appears to record whether businesses *sponsor* a plan, without regard of how long the plan has been in place. Indeed, nothing is reported on the age of the plan or whether an adviser was involved when the plan was set up.

Second, the survey asked respondents "to select all of the advisors that they consult in the management of their business" (Oliver Wyman 2015, p. 12). Given the focus of the study, of course advisers are relevant only if they were involved with the company's retirement plan. However, the study did not restrict advisers to those who provided assistance with a retirement plan. For example, a firm may have hired a financial adviser solely to assist with succession issues or asset management, but the study would credit this adviser with setting up the firm's retirement plan.

Third, consider two phases of retirement plan formation: the decision to set up a plan and the process of setting it up. Financial advisers may or may not play a role in either phase. In some cases, financial advisers may have convinced small businesses to set up a retirement plan; in other cases, the small business may have decided to set up a retirement plan and consulted a financial adviser to guide it through the process. Given the study's focus on the role of investment advisers, it would make sense to include detailed questions in the survey about that role and about the timing of the adviser's involvement.²³ Unfortunately, Oliver Wyman's own survey did not include such questions (or the authors chose to not discuss them).

The description of the survey method states that the survey had a stratified design, and "[t]his design allowed us to isolate the impact that financial advisors have upon small businesses" (Oliver Wyman 2015, p. 40). This statement is incorrect; insofar as reported, the survey does not permit any conclusions about the causal effects of advisers on retirement plan sponsorship, and it is especially lacking with respect to the role of advisers who assisted with setting up retirement plans and who were compensated in a conflicted manner.

The Oliver Wyman study also fails to consider that the involvement of a financial adviser may be correlated with other factors that affect the rate of plan formation

²³ Insofar we are aware, neither the questionnaire nor the survey's microdata have been made available.

among small businesses. For example, Brady and Bogdan (2014) found that workforce composition appears to be a primary cause for the lower rate at which small employers sponsor retirement plans. Employees who work for firms that do not sponsor retirement plans are more likely to be younger, have lower earnings, and have less attachment to the workforce—all characteristics associated with being less focused on saving for retirement. By the same token, companies with such employees may be less likely to spend money on financial advisers than, say, high-tech start-ups with a highly educated workforce.

The Oliver Wyman study does not provide clear indications of the extent to which financial advisers deserve credit for companies' decisions to form a plan or the extent to which financial advisers helped guide small businesses through the formation process, especially since a non-trivial fraction of small businesses sponsor a plan without involvement of a financial adviser (46%; see Oliver Wyman 2015, p. 14). It appears plausible that financial advisers played a role in the formation of a number of retirement plans, but the Oliver Wyman study does not support any quantification.

II. Role of Financial Advisors in Helping Individuals Save for Retirement

The second part of the Oliver Wyman study focuses on the role of financial advisers in helping individual investors. Some of it applies to DC plan investments, some to IRA investments, and some to after-tax accounts. The analysis is based on the Oliver Wyman Retail Investor Retirement Survey 2014, a survey of non-retired individuals with investments or retirement accounts. The study notes there were 4,393 valid complete responses. The analysis also draws on data from IXI Services, reportedly representing approximately 20% of U.S. consumer invested assets on a household level and approximately 30% of U.S. consumer invested assets on an account level.

Part II starts out by demonstrating that advised individuals had more financial assets than non-advised individuals. This pattern was borne out in data from both Oliver Wyman's own investor survey and from IXI Services. The Oliver Wyman study does not discuss causality, leaving open the possibilities that advisers enrich their clients or that wealthier investors are more likely to seek advice than their less wealthy counterparts. The study merely establishes a correlation, does not discuss the direction of causality, and does not recognize that financial advisers cannot claim full credit for the greater wealth of advised individuals.

Part II continues with arguments that individuals with a financial adviser are better investors along several dimensions:

- A. *Developing and maintaining a personalized financial plan.* This section (Oliver Wyman 2015, pp. 18-23) mostly draws on external research into why households save, what they value in advisers, how commonly they use plan advice offered through their DC plan, how much they contribute to their DC plan, why they roll over DC assets into an IRA, and how common DC plans and IRAs are. None of this demonstrates (or even suggests) that advised individuals are more likely to develop and maintain a personalized financial plan than non-advised individuals. More generally, none of it compares individuals with and without a financial adviser, with a partial exception in the finding that DC plan participants who used "at least one type of support contributed an average of 2.0 percentage points more of their salary to a DC plan (6.7% vs. 4.7%)" (Oliver Wyman 2015, p. 20). The support types alluded to here included educational materials, tools, and advice options, i.e.,

more types than just financial advisers. The study does not discuss causality, and indeed causality may go both ways: educational materials and other support types may prompt some DC plan participants to increase their contributions, and people with sizable DC plan balances may be more likely to seek support because they have more at stake than people with small balances.

- B. *Commitment to regular saving and investment.* This section shows that households with a financial adviser are more likely to own an IRA and that their average IRA balance is higher than that of non-advised households. External data show similar patterns for DC plans. Again, the study does not discuss causality, and indeed causality may go both ways: financial advisers may nudge people to save more in IRAs or DC plans, and people with large IRA or DC plan balances may decide to seek professional advice because they have more at stake than people with small balances.
- C. *Constructing and maintaining a well-diversified portfolio of appropriate investment products.* This section shows that advised households own more diversified portfolios than non-advised households. We agree that, all else equal, a well-diversified portfolio is generally preferable over a highly concentrated portfolio. However, the advice and the diversification are not free of charge, and a central issue is whether the diversification as advised by financial advisers generated long-term benefits. The Oliver Wyman study is silent on issues of cost and rates of return, but several academic studies suggest that the net contribution of certain financial advisers on portfolio performance is negative.²⁴ Indeed, several portions of the Oliver Wyman study are suggestive of expenses incurred with diversification. For example, “Non-advised individuals hold 70% more of their equities exposure in individual securities compared to advised individuals” (Oliver Wyman 2015, p. 28). Accepting for now that individual securities are indicative of a lack of diversification, investors incur expenses in the leading alternative—mutual funds—in the form of front-end sales loads and expense ratios. Also, Figures 20 and 21 show that advised individuals hold far more of their portfolios in variable annuities than non-advised individuals. Variable annuities are widely known to be subject to high fees.²⁵
- D. *Staying invested in the market.* This section shows that advised individuals hold less cash, as a fraction of their portfolio or IRA, than non-advised individuals. It further asserts that “Financial advisors help individuals avoid premature IRA distributions — 76% of heads of households that made traditional IRA withdrawals in 2013 were retired” (Oliver Wyman 2015, p. 34). The assertion appears to be based on a finding that most (88%) IRAs are held in a brokerage model, where the account holder has access to a financial adviser. However, the Oliver Wyman study does not present statistics about withdrawals by non-advised individuals and its evidence does not support the conclusion that financial advisers help avoid premature IRA distributions.
- E. *Periodically re-balancing investment holdings to restore desired asset allocation and risk levels.* This section shows that advised individual rebalance certain portions of their portfolio more frequently than non-advised

²⁴ See, for example, the studies reviewed in Council of Economic Advisers (2015).

²⁵ See, for example, Kaplan (2012) and Scism (2012).

individuals. Similar to the above discussion related to portfolio diversification, we agree that, all else equal, rebalancing is generally desirable. However, there may again be costs associated with selling assets and buying other assets, and front-end load fees in particular can erase any benefits of rebalancing. The Oliver Wyman study is silent on such costs.

The arguments that individuals with a financial adviser are better investors apparently are intended to convince the reader of the value of financial advice. However, the Proposed Rule is concerned with conflicted advice only, not with financial advice in general. The Oliver Wyman Retail Investor Retirement Survey 2014 does not define what it means with "financial adviser," how it asked the respondents whether they consulted a financial adviser, or how the adviser was compensated. Given the focus of the study and the fact that Oliver Wyman designed its own survey, it is puzzling why Oliver Wyman did not distinguish conflicted and non-conflicted advice. It appears Oliver Wyman assumed that conflicted advice is as valuable as non-conflicted advice. Also see our discussion of common themes in Section 2.

The study also does not consider whether non-advised individuals participate in other retirement vehicles such as DC plans. These individuals may place a higher emphasis on their DC assets which could explain some of the differences in account characteristics such as average IRA balance and diversification of their portfolios. For example, in a study on mutual fund ownership through investment professionals, Schrass (2013, p.8) finds that "mutual fund-owning households without advisory relationships were more likely to hold mutual funds only through employer-sponsored retirement plans".

In short, Part II of the Oliver Wyman study shows that financial advisers tend to be involved with relatively successful individual investors, but it does not address to what extent financial advisers deserve credit for that success, it does not address whether the price individual investors pay for financial advice exceeds the benefits, and it fails to single out conflicted advice.

Data Issues

As noted earlier, much of the analysis in the Oliver Wyman study is based on a survey of small businesses, a survey of individual investors, and data from IXI Services. We now discuss each data source in turn.

Oliver Wyman Small Business Survey 2014

According to the "Survey methodology" section (Oliver Wyman 2015, p. 40), the Oliver Wyman Small Business Survey 2014 is a survey of owners and HR decision makers of payroll-based businesses with between 1 and 100 employees. Among others, it formed the basis of the claim that small businesses with a financial adviser are more likely to set up a retirement plan than businesses without a financial adviser. See the study's Figure 7, transcribed here in Table 2.²⁶

²⁶ As discussed above, Oliver Wyman's use of the term "plan formation rates" is misleading; the rates refer to plan sponsorship. Also see the labeling of the study's Figure 7 ("Percent of businesses offering retirement plan").

Table 2. Plan Sponsorship Rates by Size of Firm and Adviser Status

	Number of employees			Overall
	1-9	10-49	50-100	
With a financial adviser	51%	77%	89%	69%
Without a financial adviser	26%	52%	75%	46%
Overall	36%	63%	80%	56%

Source: Oliver Wyman Small Business Retirement Survey 2014.

The plan sponsorship rates in Table 2 are substantially higher than nationwide sponsorship rates among small businesses reported elsewhere. For example, according to the Bureau of Labor Statistics (2014), 45% of establishments with 1-99 employees sponsored a retirement plan in 2014.²⁷ In contrast, the Oliver Wyman survey reported a rate of 56%. Also, according to tabulations by Brady and Bogdan (2014) and Copeland (2014), 17% of employees at firms with 1-9 employees had access to a retirement plan at work, whereas the Oliver Wyman survey suggested as many as 36% of firms with 1-9 employees sponsored a plan.²⁸ There are differences between Oliver Wyman's survey results and external sources within the report itself. For example, Figure 4, which is based on a Social Security Administration study (Dushi et al., 2011, Table 2), shows that 70% of employees at firms with 50-99 employees had access to a retirement plan, whereas Figure 7 reports that 80% of firms with 50-100 employees sponsor a retirement plan.²⁹ Such differences in a key metric call into question the validity of the Oliver Wyman Small Business Retirement Survey 2014. More generally, the Oliver Wyman study provides few details about the small business survey's design, sampling frame, questionnaire, response rate, or implementation.

Oliver Wyman Retail Investor Retirement Survey 2014

The "Survey methodology" section (Oliver Wyman 2015, p. 40) explains that the Oliver Wyman Retail Investor Retirement Survey 2014 was stratified by age, income, and the presence of a financial adviser. It does not state from what sampling frame the sample was drawn. Even though income was used for stratification, sampling weights were based on assets not income (Oliver Wyman 2015, p. 40.) The authors defend their unusual approach as follows: "Although we sampled based upon age, income and the presence of a financial advisor, we scale our sample to the population using age, assets, and the presence of a financial advisor, as the distribution of household assets is better documented in secondary sources than the distribution of personal income" (Oliver Wyman 2015, p. 40). They subsequently state that they used the Survey of Consumer Finances (SCF) to calculate sampling weights. However, the SCF contains detailed questions about individual and

²⁷ Some of these establishments belonged to a larger firm with multiple establishments. Since plan sponsorship tends to increase with firm size, the Bureau of Labor Statistics figures imply that sponsorship among firms with 1-99 employees was less than 45% in 2014.

²⁸ The unit of observation of Brady and Bogdan (2014) and Copeland (2014) was an employee and that of the Oliver Wyman survey, a firm. Since sponsorship rates tend to increase with firm size and larger firms employ more people, the employee-weighted rate is higher than the firm-weighted rate. The actual discrepancy is thus even larger than the difference between 17% and 36%.

²⁹ The actual discrepancy is again larger because Figure 4 is employee-weighted and Figure 7 is firm-weighted; see footnote 28.

household income, raising questions why Oliver Wyman chose assets instead of income to calculate sample weights.

Another issue relates to the definition of a "financial adviser." The study does not define the term for its stratification purposes or for its weighting purposes. It states only that weighting was based on the 2013 SCF. However, the SCF did not ask about "financial advisers." It did ask about sources of information used to make decisions about saving and investments. The SCF respondent could choose from a number of options, including lawyer, accountant, banker, broker, and financial planner, but "financial adviser" was not among the options. It thus remains unclear how to interpret the study's use of the term "financial adviser." The study does not even report what fraction of households in its survey consulted a financial adviser, other than "By one measure, 58% of households with under \$100,000 in investable assets, and 75% of non-retired households with over \$100,000 in investable assets, solicit professional financial advice" (Oliver Wyman 2015, p. 19 and attributed to the SCF). In our own analysis of the SCF we were unable to replicate these rates, but roughly approached them by including bankers, brokers, financial planners, dealers, and insurance agents. Perhaps these categories jointly formed the basis of the statement about advice rates, but we cannot think of a data source that could serve as the sampling frame for the survey's stratification by presence of a financial adviser in any of those categories. The Oliver Wyman Retail Investor Retirement Survey 2014 likely used another definition. However, the Oliver Wyman study provides few details about the investor survey's design, sampling frame, questionnaire, response rate, or implementation. Lack of public access to the survey and the discrepancies noted above do not enhance the credibility of the Oliver Wyman study.

IXI Services

In addition to its proprietary investor survey, Part II of the Oliver Wyman study relied on data from IXI Services. Even though it repeatedly refers to these data as household-level or account-level data (e.g., footnotes 25, 36, 40, 42, 43, 44, 47, 48, and 49), it is our understanding that Oliver Wyman did in fact not analyze account-level or household-level data from IXI Services. Instead, the data appear to have been aggregated to segment-level information: "Our analysis leveraged IXI Services data containing segment-level detail on U.S. consumer invested assets. Segments were defined by specific age tiers (five), income tiers (eleven), wealth tiers (seven), advisor relationship type (Full Service Brokerage vs. Discount Brokerage) and year" (Oliver Wyman 2015, p. 41). The same page explains that IXI data contain information on total segment assets, total segment number of households, et cetera.

Indeed, results from IXI Services data tend to be phrased in awkward and potentially misleading terms. For example, "94% of households examined belonged to an age / income / wealth segment in which advised households held $\geq 25\%$ more IRA assets compared to nonadvised households" (Oliver Wyman 2015, p. 23). Or, "72% of households belong to a segment in which advised households hold more than 20% less of their assets in equities" (Oliver Wyman 2015, p. 27). Such segment-level statements can be misleading, in part because either all or none of the households in a segment support the statement without regard to differences within segments.

5. INVESTMENT COMPANY INSTITUTE

Summary

In response to DOL's Proposed Rule, Brian Reid and David W. Blass of ICI filed a comment letter in July 2015 (ICI 2015a) and follow-up letters in September 2015 (ICI 2015b) and December 2015 (ICI 2015c). This section contains a review of ICI's comments.

ICI's comments criticize certain academic studies upon which DOL relied in estimating the impacts of the Proposed Rule. Separately, ICI presents alternative estimates of the performance of funds with front-end loads, with the primary conclusion that no-load funds outperform funds with front-end loads by an annual average of only 7 bps. Further, ICI asserts that the Proposed Rule would effectively eliminate accounts with front-end loads, resulting in increased annual costs for investors with assets over \$100,000, which in turn would reduce annual returns by 61 bps as these investors migrated to fee-based arrangements. Finally, ICI asserts that the Proposed Rule will eliminate advice for investors with accounts under \$100,000, resulting in a 300 bps reduction in annual returns for those investors. ICI's assumptions and calculations produce an estimate of increased costs to investors of \$18.8 billion in the 10th year.

Our primary conclusions are that (1) ICI's criticisms of the academic literature and front-end load performance results do not undermine DOL's estimates of the benefits from reducing conflicted advice and (2) ICI's estimates of the costs to investors of having to pay more for and/or losing financial advice are based on unsupported assumptions that are contradicted by information provided by other commenters. In particular, not only do ICI's criticisms of the academic literature fail to undermine DOL's interpretation of those studies, ICI's finding that the average annual returns for no-load funds exceed the annual returns for front-end load funds by 43 bps—the result that most closely aligns with the academic study DOL used in the RIA—is reasonably close to the estimated benefit from less conflicted advice described in the RIA.

With regard to ICI's estimates of the cost of the Proposed Rule, ICI offers no support for its assumptions that investors currently selecting front-end load funds would either have to pay as much as more active incumbent fee-based investors or lose access to advice. The first assumption ignores the likely emergence of new fee structures or products to continue to service investors that ICI characterizes as placing less demand on financial advisers. ICI's companion assumption that investors with balances under \$100,000 would no longer receive financial advice is inconsistent with the fact that a large proportion of investors with accounts at or below this level have the fee-based accounts that ICI presumes are too costly to provide.

Synopsis

ICI's comments criticize certain academic studies upon which DOL relied in estimating the impacts of the Proposed Rule. Separately, ICI presents alternative estimates of the performance of funds with front-end loads, with the primary

conclusion that no-load funds outperform funds with front-end loads by an annual average of 7 bps. Further, ICI asserts that the Proposed Rule would effectively eliminate accounts with front-end loads, resulting in increased annual costs for investors with assets over \$100,000, which in turn would reduce annual returns by 61 bps as these investors migrate to fee-based arrangements. Finally, ICI asserts that the Proposed Rule will eliminate advice for investors with accounts under \$100,000, resulting in a 300 bps reduction in annual returns for those investors. ICI's assumptions and calculations produce an estimate of increased costs to investors of \$18.8 billion in the 10th year.

Discussion

ICI's Criticisms of Academic Literature

Because DOL's quantitative estimates of the impact of the Proposed Rule rely heavily on the results presented in Christoffersen et al. (2013), ICI focuses the majority of its criticisms on this article.³⁰ These criticisms, which generally overlap with those of other commenters, include³¹ (1) the study does not measure the effect of the difference between fiduciary and broker advice,³² (2) the age of data used in the study,³³ (3) the application of the relationship between excess load (broker compensation in excess of the expected level) and fund performance to changes in the *average* load, and (4) the fact that Christoffersen et al.'s analysis was not weighted by assets or sales.³⁴

In Section 2 (Common Themes) we address each of these criticisms. With regard to the need for a direct measure of the effect of the difference between fiduciary and broker advice, while we are not aware of publicly available studies that explicitly compare the effects of fiduciary and broker advice, the academic literature upon which DOL relies addresses underperformance due to conflicts of interest. This focus is consistent with the target of DOL's Conflict of Interest Proposed Rule.

Because DOL applies a *relationship* between broker compensation and fund performance to compensation levels expected to occur under the Proposed Rule, ICI's concern about the age of the data is misplaced. ICI replicates Christoffersen et al.'s (2013) regression model with data from 2010 to 2014, which demonstrates the robustness of the relationship DOL applied in its impact analysis (ICI 2015c, p. 9):

³⁰ The authors' letter responding to ICI's criticisms concluded that none of them are valid; see Christoffersen and Evans (2015).

³¹ ICI also criticizes Christoffersen et al.'s relationship between fund inflows and broker compensation. Since DOL's calculations did not rely on this relationship, we do not address this criticism.

³² "Christoffersen et al. do not measure or test whether these returns were lower than what investors would have received had they used a fiduciary adviser" (ICI 2015a, p. 13).

³³ "The sample period in the paper extends from 1993 to 2009, relying largely on fund performance that is 10 to 20 years old" (ICI 2015a, p. 13).

³⁴ "Nor does the paper provide asset-weighted or sales-weighted returns to demonstrate how investors who use broker-sold funds perform as a group relative to those using similar funds in their Morningstar category" (ICI 2015a, p. 13).

The results in the second-stage regression are also in all their important elements very similar to those reported by CEM [Christoffersen et al. 2013] [...] We find a coefficient estimate on the residual load fee paid to unaffiliated brokers of -0.64 percent, which implies *an even larger effect than the -0.4972 coefficient reported in CEM.* (Emphasis added)

While the data used in academic studies may be dated, findings on incentive effects remain valid; see Section 2.

Contrary to ICI's claims (and those of others), DOL applied the relationship between fund performance and excess load properly. In particular, as explained in Section 2, Christoffersen et al. use their relationship the same way as DOL has to explain how performance improves when front-end loads are reduced. Whether some funds are above average and others are below average is irrelevant; the model is applicable to all funds when loads change.

ICI's specific criticism is that while Christoffersen et al.'s relationship between front-end load paid to brokers and performance is the result of a regression model that explains the annual returns of a fund in excess of its Morningstar category average by the excess front load payments,³⁵ Christoffersen et al. and the RIA apply the results to the total load paid to brokers, not the excess load. In particular, ICI (2015a, p. 15, emphasis in original) states:

When they attempt to measure the economic significance for the investor, they incorrectly multiply the coefficient of the "excess load" variable by the *average load paid*, and argue that the typical fund underperforms by 1.13 percent annually. But the regression relating fund performance and loads was not run using actual load, but using "excess load." The residuals from their first regression measuring the "excess load" should have a mean of *zero*. Taking the results from their analysis literally, they should conclude that the average broker-dealer funds neither underperform nor outperform their Morningstar category average.

DOL's application of the relationship between excess loads and fund performance is valid because a reduction in load would improve a fund's performance regardless of whether the fund in question paid brokers an above-average amount (in which case excess load would be positive) or a below-average amount (in which case excess load would be negative). Accordingly, ICI's (and NERA's) criticism is invalid. While it is true that residuals have a zero mean,³⁶ Christoffersen et al. and the RIA are not using the model to explain the average effect over all funds used to estimate the model, but rather the effect of an overall *change* in excess load.³⁷

³⁵ The excess load variable is calculated as the actual payment to brokers minus payment predicted by a regression model that explains payments to brokers as a function of fund characteristics and whether the broker is captive or unaffiliated.

³⁶ Because the number of observations (163,347) in the regression model that produced the residuals (Christoffersen et al. 2013, p. 217) is somewhat larger than the number of observations (113,153) in the regression model explaining performance (Christoffersen et al. 2013, p. 226), the average of the residual used as an independent variable in the latter would not necessarily be zero.

³⁷ ICI reiterates its criticism in its December 2015 letter to DOL:

Finally, while weighting often makes sense when calculating averages, Christoffersen et al. did not calculate such an average, but instead developed the relationship between excess load and fund performance. As we discussed earlier, there is no requirement in econometric theory or practice for the observations used to establish such a relationship be weighted when they differ in size by some measure.³⁸ In fact, in its latest comment (ICI 2015c, p. 3) now agrees that there is no problem with the Christoffersen et al. relationship:

The Department needed to weight the research findings taken from the CEM study by assets or sales of fund shares. This problem remains *even though* the CEM study properly adjusted for the levels of funds' assets in its regressions. The CEM study, like most of the other academic studies the RIA cites, conducts its analysis at the fund level. The RIA seeks to convert this fund-level analysis into aggregate dollar effects on the total IRA assets invested through broker-sold funds. To translate fund-level findings into *market-level* dollar effects, the Department would need to weight the fund-by-fund effects predicted by the CEM regression by the asset levels or sales of those funds. The RIA did not do that.

ICI's claim is incorrect. Because DOL applied the relationship to the *change* in excess load expected from the Proposed Rule and that change is asset-weighted, DOL's calculation is a proper application of Christoffersen et al.'s regression relationship.³⁹

The Department did in fact misapply a regression coefficient taken from the CEM study. To be clear, this has very little to do with the results in the CEM study, which stand on their own merits [...] The Department erred by applying the CEM coefficient to the front-load paid to brokers rather than to the residual load paid to brokers, inflating the Department's estimate of the benefit of its proposed regulation. (ICI 2015c, p. 3.)

As a preliminary matter, DOL applies Christoffersen et al.'s results the same way the authors themselves apply the relationship between excess load and performance (Christoffersen et al., 2013, p. 228). More fundamentally, the regression relationship is properly applied to the *change* in the variable affected by the policy and not the current level of the variable as ICI suggests. ICI's misguided approach is analogous to a study which first centered the data, estimated a relationship between a key independent variable and the dependent variable, and then insisted that the coefficient be used with the mean of the key variable (which would be zero by construction) and from this exercise concluding that there was no effect.

³⁸ Christoffersen and Evans (2015, p. 2) provide an explanation why their econometric approach did not require asset weighting.

³⁹ To see why, note that the relationship is being applied to the change in excess load. Therefore, at the individual fund level, the model would produce the following: $\Delta \text{return}_i = \beta \Delta \text{load}_i$, where β is the regression coefficient, Δreturn_i is the predicted change in return for fund i , and Δload_i is the change in excess load for fund i expected from the Proposed Rule. The overall change in return is obtained by weighting the predictions for each fund, and then summing them.

$$\Delta \text{return} = \sum w_i \Delta \text{return}_i = \sum w_i (\beta \Delta \text{load}_i) = \beta \sum w_i \Delta \text{load}_i = \beta \Delta \text{load}.$$

ICI's Analysis of Front-End Load Fund Performance

ICI reports average returns of front-end mutual funds in a number of ways. The results, which are based on Morningstar data, include returns for domestic equity, international equity, taxable bond, and balanced funds. ICI does not provide the specific data used to produce the results (to the extent those data differ from data readily available to other users), nor does it provide details often available in academic articles, such as the definitions of the fund categories and the number of observations used to calculate average returns.

ICI's calculations start with annual returns, net of expenses, for 2008-2014. ICI describes the calculation as follows (ICI 2015a, p. 16):

To measure the experience in broker-sold share classes, we use gross sales and assets of front-end load share classes from 2007 through 2013 and measure the performance of these share classes or their funds in subsequent years to capture what investors would have experienced if they stayed in their funds. The reason for focusing on the more recent time period is that the mutual fund market has changed significantly in the past twenty years, as we discussed in Section II. We then calculate fund returns, net of fund fees, based on Morningstar data.

As a baseline, we take one-year net returns of share classes with front-end loads from 2008 through 2014 and subtract each share class's Morningstar category return from the same year to create a relative return.⁴⁰ To measure how investors as a group using front-end share classes perform, we then weight each fund's relative performance in the subsequent one-year period by sales or assets from the reference year. Similar measures are used for retail no-load funds to provide a basis for comparison.

Table 3 lists ICI's return results.

**Table 3. Front-End and No-Load Fund Returns Calculated by ICI
(Annual Percent Relative to Morningstar Category Average)**

Measure	Front-End		Difference	Source
	Load	No Load		
Simple Average	0.13	NA	NA	Figure 2
Sales Weighted	0.27	0.70	0.43	Figure 4
Average of yearly returns_sales weighted	0.16	NA	NA	Figure 3
Average of yearly returns_asset weighted	0.15	NA	NA	Figure 3
Average of yearly 3-year returns_sales weighted	0.17	0.44	0.27	Figure 5
Average of yearly 3-year returns_asset weighted	0.37	0.65	0.28	Figure 5

Table 3 demonstrates the following.

- ICI observed that the average net return, relative to Morningstar categories, was 27 bps with a sales-weighted average and 13 bps with a simple

⁴⁰ ICI (ICI 2015b, p. 3) characterized its approach as being the same as that used by Christoffersen et al. (2013) and other academic studies.

average.⁴¹ ICI (2015a, p. 17) credits brokers for this outcome: "The fact that the sales-weighted average exceeds the simple average suggests that brokers tended to guide their clients to funds that subsequently slightly outperformed, not underperformed."⁴²

- The approach that most closely aligns with DOL's use in the RIA of Christoffersen et al.'s (2013) findings is the sales-weighted approach shown in the highlighted, second row. ICI finds that load funds underperform no-load funds by 43 bps. To put this difference into perspective, the RIA (DOL 2015, p. 115) estimates that reducing conflicted advice would increase annual returns by about 50 bps in the latter years of the 2017-2026 period (first scenario).
- ICI reports averages of one-year returns for 2007 through 2013 on sales-weighted and asset weighted bases. These averages are lower than the overall average reported in the second row (0.16 percent or 0.15 percent versus 0.27 percent), with only a minimal difference between the sales-weighted and asset-weighted results.
- ICI also reports the averages of three-year returns for the period 2007-2011.⁴³ The superior performance of the no-load funds is 27 or 28 bps for these comparisons. In contrast to the minimal difference in one-year returns listed in the third and fourth rows, ICI's asset-weighted three-year returns are curiously about 20 bps higher than the corresponding sales-weighted returns.

In addition to reducing the performance gap between no-load and front-end load funds by introducing three-year returns, ICI (2015a, p. 21) adds back 12b-1 fees, which reduces the performance gap by an additional 20 bps. Christoffersen et al. (2013) did *not* make this adjustment. ICI then uses the resulting gap of 7 bps to quantify the net costs it attributes to the Proposed Rule, which we discuss in the following sections.

The appropriateness of adding 12b-1 fees to investment returns is debatable. Presumably, the argument is that they serve to compensate brokers for their services, just like fees do in fee-based accounts. In a recent paper that was also cited by ICI (2015c), Reuter (2015, p. 6) observes that adding back 12b-1 fees "is reasonable except to the extent that conflicts of interest lead brokers to recommend funds that charge higher 12b-1 fees in order to pay higher commissions." Also, one could argue that the broker was already compensated through a share of the front-end load at the time of purchase; it is unclear whether investors are aware that they continue to pay the broker for as long as they own the fund and whether they would consider the 12b-1 fees as part of their rate of return.

⁴¹ The average for sales-weighted front-end load funds listed in the second row differs from the corresponding average in the third row because the former is a single average for the entire period, while the latter is the average of the annual averages.

⁴² Whether brokers encouraged investors to select better-performing front-end load funds is a different issue from whether brokers advised investors to select front-end load funds that underperformed alternative funds, such as no-load funds.

⁴³ The end of the three year period for calculating a three-year return for funds sold in 2011 is 2014—the last year of the data used by ICI.

In summary, rather than demonstrating errors in Christoffersen et al. that cause the results in the RIA to “collapse” (ICI 2015a, p. 5), ICI’s result that most closely aligns with Christoffersen et al.’s approach—the 43 bps superior performance of no-load funds as measured by one-year net returns—is quite similar to the approximately 50 bps impact in the RIA’s first scenario.⁴⁴ The narrower gap claimed by ICI required changes—the use of three-year returns and the adding back 12b-1 fees—from the measurements used in the Christoffersen et al. study.⁴⁵

ICI’s Claims about Increased Costs for Larger Investors

ICI (2015a, p. 25) also claims that the Proposed Rule will result in fewer investors being able to select commission-based funds:

[T]he BIC exemption is unworkable; even if could work, it would impose prohibitive costs on brokers. Brokers subject to the Exemption’s many limitations, burdens, and costs, as well as its increased exposure to liability, are likely to seek to move many of their clients to fee-based accounts. Such accounts, however, require a much greater level of time and engagement through frequent rebalancing of investors’ accounts a level of service that is unnecessary for an investors with a modest balance who is typically better off as a buy-and-hold investor. This additional ongoing engagement results in higher and ongoing expense to the investor.

As explained in more detail below, ICI assumes that the shift from commission-based to fee-based accounts would eventually reduce annual returns to investors by 61 bps, which would exceed the 7 bps gain from reduced conflicted advice by 54 bps. As described in the next section, ICI also assumes that investors with balances under \$100,000, which according to ICI account for 19 percent of current front-load IRA funds, would no longer receive any financial advice; therefore, ICI assumes that investors accounting for 81 percent of assets in traditional IRAs would be shifted to fee-based accounts. By the tenth year, by which time ICI’s calculations assume that account balances that existed before the rule would have fully turned over,⁴⁶ ICI

⁴⁴ ICI’s result is also very similar to Reuter’s (2015, p. 13) conclusion:

Within the broader sample of actively managed funds [...] the difference is 0.47% if I focus on category-adjusted after-fee returns and 0.20% [...] if I add back 12b-1 fees. To the extent that conflicts of interest lead brokers to recommend funds with higher-than-average 12b-1 fees (as performance differences between active and passive broker-sold funds suggest), the actual performance difference within the broader sample of actively managed funds is likely to fall between 0.47% and 0.20%.

⁴⁵ There are additional methodological differences that would need to be considered in order to explore fully the differences between ICI’s and Christoffersen et al.’s results. For example, Christoffersen et al. (2013, page 226) controlled for a number of factors, such as fund size, inflows, and redemptions, while ICI’s analysis reports no such controls. Not accounting for such factors could cause results such as averages or weighted averages to be biased.

⁴⁶ ICI uses the asset turnover distribution from the RIA (DOL 2015, p. 114, Table 3.4.1-2), which posits that 16.8 percent of assets turn over in the first year, with declining percentages in subsequent years so that all assets have turned over by the 10th year.

estimates that the net cost to investors who were shifted to fee-based accounts would be \$8.2 billion.⁴⁷

ICI provides no analysis or quantitative estimates of how many accounts would be moved from commission-based to fee-based and the level of assets associated with such moves. Instead, ICI simply assumes that *all* investors in commission-based accounts would either be moved to fee-based accounts and pay more in fees as a result or have balances not sufficient for advisers being willing to service the accounts. As discussed in Section 2, this presumption is contradicted by evidence from other countries.

ICI's (2015a) calculations ignore that a portion of IRAs is held in discount brokerage accounts. These account holders do not receive advice, and their accounts are not affected by the Proposed Rule. ICI's (2015a) calculations of the share of accounts and assets that will migrate or lose access to advice are thus overstated.

For those investors whom ICI assumes will be shifted to fee based accounts, ICI assumes that costs would increase by the *average* difference of 61 bps between expenses for fee-based (average of 111 bps) and commission-based (average of 50 bps) accounts. In other words, rather than account for the lower demands incumbent commission-based investors impose by introducing new fee structures or new products, ICI casually assumes that advisers will charge them fees based on services that they do not demand, such as frequent rebalancing. ICI did not discuss the possibility that because of their lower trading volumes and account turnover, current commission-based investors who had to migrate to fee-based accounts (or perhaps some other mechanism) would probably not be as costly to serve as incumbent fee-based investors.

ICI's Claims about Loss of Advice for Smaller Investors

ICI (2015a, p. 27) further assumes that investors with smaller account balances would completely lose financial advice:

[F]ee-based accounts may not be available to low- and middle-income IRA investors who cannot meet minimum account balance requirements. Currently, fee-based advisers often require minimum account balances of \$100,000 [...]

ICI (2015a, pp. A-1 to A-2) further assumes that investors losing advice will eventually experience annual returns that are 3% lower than the returns they earned in front-end load accounts:

We assume that these assets underperform by 3 percent a year compared to their performance with a broker [...] The 3 percent underperformance reflects lower allocation to stocks and higher allocation to cash, early withdrawals and elimination of tax deferral, and poor market timing decisions.

ICI's calculates the impact of lost advice by combining the two assumptions—investors holding 19% of assets experience a 3% loss in annual return. By the tenth year, ICI estimates that the net cost to investors who lost investment advice would

⁴⁷ ICI (page A-3) adopts DOL's projection of \$1.868 trillion in assets by the 10th year. Therefore, ICI's estimate of \$8.2 billion = 1,868 x 0.0054 x 0.81.

be \$10.6 billion.⁴⁸ Combined with the loss from investors with accounts in excess of \$100,000 discussed in the previous section, ICI (2015a, p. 30) reports a total loss of \$18.8 billion by the tenth year.⁴⁹

ICI provides no support for either of its assumptions, both of which are inconsistent with information provided by other commenters. With regard to the loss of advice, the results of NERA's database of over 63,000 accounts show many fee-based accounts with balances well under \$100,000 (NERA 2015a and 2015b). Table 4, constructed from NERA's results, displays this information.

Table 4. Fee-based and Commission-based Accounts by Account Size

Balance	Fraction of fee-based accounts	Fraction of commission-based accounts	Percentage of accounts that are fee-based
\$1K-\$10K	2%	23%	3.5%
\$1K-\$25K	9%	41%	8.4%
\$1K-\$50K	22%	57%	13.8%
\$1K-\$100K	42%	72%	19.5%
\$1K-\$250K	72%	87%	25.6%
\$1K-\$1M	97%	98%	29.1%
\$1K+	100%	100%	29.4%

Source: NERA (2015a, 2015b).

The first two columns display the cumulative distributions of fee-based and commission-based accounts by account size. For example, 2% of fee-based accounts have balances from \$1,000 to \$10,000 and 42% have balances of \$100,000 or less. ICI's assumption that accounts with balances less than \$100,000 would be too costly to serve cannot be reconciled with the fact that a large proportion of fee-based accounts have balances below \$100,000. The last column of the table shows the proportion of accounts that are fee-based. In particular, 19.5 percent of accounts in NERA's database with balances of \$100,000 or less are fee-based accounts.

Since ICI does not indicate how it determined that lost advice would reduce annual returns by 3%, it is not possible to evaluate how ICI reached this conclusion. For example, unlike Litan and Singer (2015a), and the Vanguard (2014a) document upon which they rely, ICI provides no detail on factors such as the specific losses that stem from lost advice (e.g., better rebalancing increases returns by x percent) or on the proportions of investors currently relying on such advice (but presumably no longer would be able to do so). Further, ICI's estimate of a 3% loss far exceeds the flawed and inflated 44.5 bps loss Litan and Singer (2015a) attribute to lost advice; see Section 7. Finally, experiences from other countries suggest that banning conflicted advice may in fact not reduce access to advice (Europe Economics, 2014).

⁴⁸ ICI (2015a, page A-3) adopts DOL's projection of \$1.868 trillion in assets by the 10th year. Therefore, ICI's estimate of \$10.6 billion = 1,868 x 0.03 x 0.19.

⁴⁹ ICI's estimate excludes the benefit from less conflicted advice for investors with accounts under \$100,000.

6. COMPASS LEXECON

Summary

In response to DOL's Proposed Rule, Compass Lexecon wrote a comment titled "Tax Consequences to Investors Resulting from Proposed Rules Relating to Financial Representative Fiduciary Status" (Compass Lexecon 2015). This section contains a review of the Compass Lexecon comment.

The Compass Lexecon comment states that as a result of the Proposed Rule, commission-based IRA accounts with balances under \$25,000 will lose access to advice. If, in order to preserve access to investment advice, investors opt to use a taxable savings account instead, these investors may experience a reduction in retirement savings. The comment focuses on quantifying the effect of this hypothetical migration to taxable savings accounts. We agree that *if* investors use taxable savings accounts instead of IRAs to fund their retirement, then they may experience reduced savings. However, we disagree with Compass Lexecon on the extent to which retirement investors will forego tax-sheltered accounts.

Among the households that Compass Lexecon identifies as at-risk, some households are presumably already in a commission-based advice relationship and some are not. Among investors already in a commission-based relationship, there may be a small subset of investors who want to preserve their *existing* relationship with their adviser and the advisory firm at all costs. That is a theoretical possibility and it can be accomplished using the Proposed Rule's carve-outs and exemptions. But given the presence of numerous comparable alternatives, we estimate this subset to be overstated in the Compass Lexecon comment. Among new investors who do not have an existing relationship, the desire to create a new relationship with an adviser using a taxable savings mechanism is expected to be even smaller.

The Compass Lexecon comment provides little to no analysis of the likelihood of investors switching to taxable savings accounts. Instead, in quantifying potential losses from investors using taxable, rather than tax-deferred accounts, it simply assumes that the bottom half of investors using brokerage accounts would use taxable accounts instead.⁵⁰ Given the current options available to investors as well as alternatives that are being introduced in the fast evolving market for investment advice, we do not foresee a consequential number of investors making this switch to taxable accounts. Current and future investors with small asset balances who seek access to investment advice already have many options to choose from. Industry trends suggest even more options may become available to them.

Also, Compass Lexecon erroneously assumes that IRAs that start small (under \$25,000) will grow to be average at the time of retirement. Instead, IRAs that start relatively small are likely to remain relatively small.

⁵⁰ Compass Lexecon (2015, pp. 19-20). More precisely, Compass Lexecon assumes that one-half of IRAs in brokerage accounts would not have had \$25,000 when the accounts were opened.

In short, while we agree that tax-sheltering can be beneficial to IRA investors, the Compass Lexecon comment relies on unsupported or incorrect assumptions on investor behavior to make overly dire predictions on eroded retirement savings.

Synopsis

The basic premise of the Compass Lexecon comment is that the Proposed Rule may cause an investor who would have opened an IRA to instead open a taxable savings account in order to preserve access to a commission-based account and the assistance that comes with it. Specifically, it expects firms currently offering commission-based IRAs will no longer find it cost-effective to offer IRAs to small account holders, such as those with a balance below \$25,000. However, firms could still offer brokerage-based taxable accounts. If investors opt for a taxable account in order to gain or retain access to advice, they will lose tax deferral benefits and end up with fewer retirement savings that will not fund their retirement for as long as investors utilizing an IRA can expect.

Section I of the comment provides background and summarizes the rationale, approach and findings of the analysis conducted. In Section II, the comment describes the model used to analyze the reduction in retirement savings, the assumptions made and other parameters used to develop the model. Section III describes how the model operates, compares different investment mechanisms (taxable savings account versus a traditional IRA and a Roth IRA), and explains the simulation techniques used to evaluate the uncertainty embedded in the assumptions. Section IV presents the results and potential impact on investors at various age and income levels, the implications for retirement security and an estimate of total potential investor losses due the Proposed Rule.

The Compass Lexecon comment finds that an investor who uses a taxable account to generate retirement savings can pay a median effective average tax rate of 30.0%-43.3% (Exhibits A and B, Median Values) compared to 17.1%-25.0% and 15.0%-25.0% for investors using Roth IRA and Traditional IRAs respectively. The ranges in these estimates are driven by uncertainty in the assumptions made—the investor's age, income, tax rates, asset allocation, returns, size and frequency of contributions, and age at retirement. The Compass Lexecon comment concludes that about 7.0 million household accounts could be affected and estimates the potential investor losses at between \$147 billion and \$372 billion over what we understand to be a period of roughly 30 years.

Discussion

Overview

Compass Lexecon uses a model to show that IRA investors can suffer a reduction in retirement savings as a result of the Proposed Rule. This effort hinges on the following line of thinking. The DOL's Proposed Rule will cause investors who use a commission-based IRA account to move to a fee-based account. But "participants in this rulemaking have stated that, if subjected to the changes in fiduciary status imposed by the proposed amendments, firms currently offering commission-based IRAs will no longer find it cost-effective to offer IRAs to small account holders, such as those with a balance below \$25,000" (Compass Lexecon 2015, p. 1). However, taxable savings plans, which are not affected by the Proposed Rule, will be available

to these investors as an alternative mechanism to fund their retirement. If an investor, unable to gain access to a commission-based IRA account and the assistance that comes with one, instead chooses to start a taxable savings account, he or she stands to lose a large portion of retirement savings to taxes every year.

Compass Lexecon attempts to quantify the impact on retirement savings, *if* investors use taxable savings accounts instead of IRA accounts. Little to no attention is devoted to how likely investors are to use taxable accounts to fund retirement. Whether the severity of the problem that Compass Lexecon seeks to highlight equates to a mountain or a molehill depends heavily on the propensity of investors to start using taxable accounts to fund retirement savings.

The remainder of this section focuses on the likelihood of investors, who by assumption cannot avail themselves of a commission-based IRA account, would instead choose to open a taxable savings account.

Likelihood of Investors Using Taxable Savings Accounts to Save for Retirement

The Compass Lexecon comment operates under the premise that some brokerage investors may consider using a taxable savings account to fund their retirement: “the proposed amendments have the potential to affect all households that (absent the amendments) would have started brokerage IRAs either from a contribution or a rollover of less than \$25,000” (Compass Lexecon 2015, p. 19.).

But the comment does not consider current and future options available to IRA investors who want to gain or preserve access to an adviser account arrangement. We have discussed multiple options available to IRA investors seeking to preserve access to financial advice such as ‘robo’ advice, target-date mutual funds and hybrid investment advice that combines automated and human-based investment advice. These options are discussed in detail in Section 2.

The presence of existing options for investment advice as well as the possibility of new options suggest that new and existing IRA investors are not likely to sacrifice valuable tax benefits to preserve access to human-based investment advice.

Benefits of Tax Sheltering

To demonstrate and measure the effect of tax sheltering, Compass Lexecon developed a model that used multiple inputs and made numerous assumptions to conclude that investors in taxable savings accounts would experience lower savings at retirement. To address uncertainty in the assumptions, Compass Lexecon evaluated multiple scenarios to estimate that “at the time of retirement, taxable saving accounts have a value that is between 11.1 percent and 21.9 percent lower than Roth IRAs, and between 18.2 percent and 28.1 percent lower than traditional IRAs” (Compass Lexecon 2015, pp. 19-20). This reduction was applied to an estimated \$1,323 billion in IRA savings to arrive at potential investor losses ranging from \$147 billion to \$372 billion by the time investors reach age 65.

We agree that investors in taxable savings accounts will experience reduced rates of savings. However, the estimated \$1,323 billion in IRA savings at retirement is inflated.

To arrive at this estimate, Compass Lexecon assumes that half of the estimated 14.0 million IRAs currently in a brokerage setting would not have \$25,000 when opening an IRA, but would open a taxable account instead. These accounts are then assumed to grow and reach the overall IRA average at age 65 of \$188,976 (in 2013 dollars). But because these affected accounts had low assets (less than \$25,000) when they were assumed to start, they are unlikely to reach the IRA average at retirement. The assumption that accounts with low balances will somehow reach average account balances at retirement is unreasonable and inflates the measure of potential effect on savings. Moreover, the total potential investor losses that Compass Lexecon calculates are accrued over 30 to 40 years of investment and have to be divided appropriately to arrive at an annual measure.

The reduced savings between IRAs and taxable accounts is also driven by the tax rates paid by investors during retirement. Tax rates are applied on anticipated retirement income. Compass Lexecon assumes that investors will experience a reduction in income of 40% upon retirement. This assumption is based on findings from a 1997 working paper and a 2008 publication from the Social Security Administration (Biggs and Springstead 2008). Using the latter source, Compass Lexecon states that the replacement rate, expressed as retirement income as a percentage of preretirement earnings, is 69% and 52% for median households in the 3rd and 4th highest lifetime earnings quintiles respectively. However, these statistics measure the replacement rate from income from shared Social Security benefits only, rather than from total household retirement income from all sources. If these additional sources of income are accounted for, it will increase Compass Lexecon's income and tax rate assumptions for IRA investors and thus reduce the benefits of tax sheltering.

7. LITAN AND SINGER

Summary

In response to DOL's Proposed Rule, the Capital Group submitted a report by Robert Litan and Hal Singer of Economists, Inc. titled "Good Intentions Gone Wrong: The Yet-To-Be Recognized Costs of the Department of Labor's Proposed Fiduciary Rule" (Litan and Singer, 2015a). In response to questions from the DOL, Litan and Singer provided additional details in a letter (Litan and Singer, 2015b). This section contains a review of the Litan and Singer study and letter.

Litan and Singer assert that DOL's Proposed Rule would result in a reduction in financial advisory services, particularly for individuals with modest investment portfolios, and in cost increases for other investors who migrate from brokerage to advisory accounts. They further assert that the requirements of the Best Interest Contract Exemption are so onerous that "it is unlikely that many brokers will seek an exemption." Consequently, the study claims that some investors would be left without financial advice, which would result in poorer financial decisions.

Litan and Singer do not provide, or cite, empirical analysis supporting their premises. The financial industry, renowned for its ability to innovate and evolve, is likely to adapt to new regulation through modified account types. Low-cost "robo" advice options, which are especially suitable for small accounts that do not need much advice, are already increasingly available, including for very small accounts. Also, investors may turn elsewhere for advice. Even apart from corrections discussed below, Litan and Singer's study could be viewed as a "what-if" exercise based on unsupported assumptions.

The study faults DOL for not including the impacts of reduced financial advice and proceeds to estimate that the loss of financial advice would reduce the annual returns of investors with modest portfolios by 44.5 bps and increase the costs of investors migrating to advisory accounts by 31 bps. These estimates of the "yet-to-be recognized costs" exceed Litan and Singer's 25 bps restatement of DOL's estimate of the Proposed Rule's benefits of reducing conflicted advice. The study also claims that (1) DOL's application of results from academic studies in estimating the gains from less conflicted advice substantially overstates the gains, (2) a simple disclosure statement would be a more cost-effective alternative for reducing conflicted advice, and (3) DOL has not produced real-world empirical support for its rejection of greater disclosure requirements.

DOL's Regulatory Impact Analysis (RIA) predicts that the Proposed Rule can generate approximately \$40 billion over 10 years in additional investment returns. Litan and Singer restate this amount as what they claim to be an equivalent annual return increase of 25 bps. In performing this translation, Litan and Singer incorrectly divide the *discounted* 10-year benefit by an *undiscounted* asset base, which has the effect of understating the benefit. Correcting this error lifts the equivalent annual return boost from reduced conflicts of interest estimated in the RIA from 25 bps to 36 bps.

Litan and Singer heavily rely on what appears to be a Vanguard training or marketing document to estimate the 44.5 bps loss in annual return they attribute to reduced financial advice. The validity of this estimate depends on (1) whether

Vanguard's results—which are based on non-conflicted advice—apply to conflicted advice, (2) whether there are costs associated with financial advice that are not accounted for in Vanguard's results, (3) whether there is double-counting among separate components of purported negative impacts, and (4) small investors' proportion of the asset base of IRA investments in funds with front-end loads (upon which the RIA's estimated impacts are based). Even if the Vanguard results are accepted as valid and applicable, correcting double-counting and other errors and weighting by the proportion of assets held by investors with modest portfolios would reduce the estimated benefit of financial advice from 44.5 bps to 2 to 3 bps. Even that effect assumes that the value that Vanguard attributes to its non-conflicted advice applies equally to conflicted advice. Similarly, Litan and Singer's estimate of a 31 bps cost for investors migrating to advisory accounts is overstated, as it relies on the flawed Oliver Wyman (2011) study that excluded costs and did not account for the fact that current brokerage investors tend to be less costly to serve. Also see Section 2. Table 5 summarizes Litan and Singer's estimates of gains and losses from the Proposed Rule and their corrected values. Instead of a net loss of 8 bps as projected by Litan and Singer, our corrections suggest a net gain of 33 bps.

Table 5. Summary of Litan and Singer Estimates and Their Corrected Values

	Litan and Singer		Corrected	
	Estimate (bps)	Asset- weighted* (bps)	Estimate (bps)	Asset- weighted* (bps)
Benefit from reduced conflicts of interest (all IRAs)	25	25	36	36
Loss from increased market timing and less portfolio rebalancing (modest IRAs)	-44.5	-6.675	-17.5	-2.625
Higher fees in advisory accounts (high-balance IRAs)	-31	-26.35	0	0
Net gain from Proposed Rule		-8.025		33.375

*Assumes that 10%-20% of assets are in IRAs with modest balances and the remainder in high-balance IRAs; see the text. This table applies a weight of 15% for modest- and 85% for high-balance IRAs.

Litan and Singer also challenge DOL's calculations of the cost of conflicted advice, asserting that the DOL misapplied or misinterpreted academic studies. Litan and Singer offer no empirical support for the validity or magnitude of their specific criticisms, which are generally undermined by a careful reading of the academic literature upon which DOL relied.

Finally, Litan and Singer attempt to support their alternative disclosure proposal with an academic article that deals with factors that mitigate, but not necessarily eliminate, the harmful effects of conflicted advice. Based upon a review of that article, we conclude that it does not support that Litan and Singer's proposed alternative would eliminate the effects of conflicted advice. In fact, the authors of that article explicitly argue for decreasing conflicts of interest rather than disclosing them.

Synopsis

Litan and Singer assert that (1) DOL's Proposed Rule would result in a reduction in financial advisory services for individuals with modest investment portfolios and cost increases for investors migrating from brokerage to advisory accounts and (2) the requirements of the Best Interest Contract Exemptions are so onerous that "it is unlikely that many brokers will seek an exemption." Consequently, the study claims that many investors would be left without financial advice, which would result in poorer financial decisions. The study faults the DOL for not including the impacts of reduced financial advice and proceeds with its own quantitative estimates of the monetary impact of a loss of financial advice. The study claims that these "yet-to-be recognized costs" exceed the benefits from reducing conflicted advice estimated by DOL.

Litan and Singer's quantitative estimates include the following:

- Based on the premise that many investors would lose access to advice and on Vanguard (2014a), Litan and Singer produce estimates of the loss in benefits from financial advice. The study calculated an impact of 27 bps for advising investors to avoid market timing and 17.5 bps for more portfolio rebalancing, for a total impact of 44.5 bps.
- Based on a single, hypothetical example presented by Oliver Wyman (2011, p. 23), Litan and Singer apply an annual estimate of a 31 bps increased cost to all investors who would migrate to advisory accounts.
- Scenario 1 of DOL's RIA predicts that the Proposed Rule will generate approximately \$40 billion over 10 years in additional investment returns. Litan and Singer convert this benefit into an annual percentage by subtracting \$240 million in annual compliance costs and dividing over the average investment base of \$1.487 trillion that the study calculates from data presented in the RIA.
- Because the resulting gain of about 25 bps from reduced conflicted advice is less than the 44.5 bps lost from reduced financial advice and the 31 bps increase in costs from migrating to advisory accounts, the study concludes that the costs of the Proposed Rule exceed the benefits.

The study also claims that (1) DOL's application of results from academic studies in estimating the gains from less conflicted advice substantially overstates the gains, (2) a simple disclosure statement would be a more cost-effective alternative for reducing conflicted advice, and (3) DOL has not produced real world empirical support for its rejection of greater disclosure requirements.

Discussion

Litan and Singer's Primary Estimates

Litan and Singer argue that the Proposed Rule would cause financial advisers to provide less advice, particularly to investors with smaller balances. A major focus of their study is a comparison of the magnitude of the benefits from reducing conflicted advice presented in DOL's RIA with the benefits that according to Litan and Singer would be foregone due to reduced advice. Litan and Singer followed these steps in carrying out the comparison: (1) translate the 10-year gains from DOL's first scenario (\$39.8 billion; RIA, Table 3.3.1-1) into an average increase in annual return

on investment and (2) use estimates from what appears to be a Vanguard training or marketing document (Vanguard 2014a) to produce estimates of what Litan and Singer offer as the loss in annual return from investors receiving less advice with regard to market timing and portfolio rebalancing. The study estimates that loss of advice would result in a reduction in annual returns that exceeds the gain that the Litan and Singer's translation of DOL's first scenario's benefits produced.

Translating DOL's 10-Year Gains into an Increase in Average Annual Return

Litan and Singer's calculation (1) starts with DOL's 10-year gain of \$39.8 billion; (2) divides this amount by 10 to produce an average annual gain that the study rounds to \$4 billion; (3) reduces this amount by DOL's estimated annual implementation costs of \$0.24 billion, producing annual net benefits of \$3.76 billion; and (4) divides the average annual net benefits by the study's estimate of \$1,487 billion for the average 10-year investment base, resulting in a gain of 25 bps.⁵¹

Because the numerator of their calculation starts with *discounted* 10-year benefits,⁵² but the denominator—average asset base—is stated in nominal dollars, Litan and Singer's translation of DOL's 10-year impact into a basis-point equivalent understates the effect on average annual return. One way to correct the study's improper mixing of real dollars in the numerator with nominal dollars in the denominator is the following calculation: (1) for each of the 10 years in DOL's 10-year scenario, calculate an annual increase in return as the change in asset differential (row F) less implementation cost of \$0.24 billion divided by the average of the beginning and ending assets (rows C and E) and (2) calculate the 10-year average of these increases. The result of this calculation is a gain of 36 bps.⁵³

⁵¹ $(4.00 - 0.24) / 1,487 = 0.25\%$. The study, which does not provide details on how the 10-year average investment base was calculated, appears to have used data in Table 3.4.2-1 of the RIA. We calculated an average investment base of 1,496 billion, using rows (B) and (D)—beginning- and end-of-year baseline front-end load mutual fund assets. Adding to this uncertainty, the study estimates the investment base to be \$1.478 trillion on page 1 and \$1.487 trillion elsewhere.

⁵² The RIA describes the calculation of a discounted (or real) 10-year gain as follows (DOL 2015a, p. 117):

The asset differential at the end of the 10-year period (2025, Row H) together with the portion of the asset differential withdrawn in each year (Row G) makes up the 10-year quantified subset of IRA investors' expected gains under alternative scenarios 1. However, before those numbers are summed, they are each discounted by the appropriate number of years at a rate of 5.3 percent (Rows I and J) so that the 10-year front-load-mutual-fund-gain-to-investors is expressed in January 1, 2016 dollars.

⁵³ The annual returns we calculate appear to match those described by Litan and Singer (p. 7): "Table 3.4.1-1 of the RIA suggests that its calculated improved performance differential, which starts out at 10 basis points, eventually will grow to 51 basis points in 10 years, as currently held IRA and defined contribution funds move to better performing funds." Litan and Singer are most likely describing Table 3.4.2-1, not Table 3.4.1-1. We matched the beginning and ending values of 10 and 51 bps. Unlike Litan and Singer's calculation of a 25 bps effect, these annual returns (as well as the average of the annual returns) do not suffer from the bias due to mixing real and nominal dollars.

Losses Due to Reduced Financial Advice

Litan and Singer's estimate that the loss of financial advice would result in a 44.5 bps reduction in annual return consists of two components: a 27 bps loss due to market timing and a 17.5 bps loss due to portfolio rebalancing. Both components are back-of-the-envelope estimates, as described next.

For the market timing estimate, Litan and Singer rely on a Vanguard comparison of the performance of self-directed investors (for which Vanguard and Litan and Singer assumed there was no advice with regard to market timing) with performance of Vanguard's Target Retirement Funds over the five years ending on December 31, 2012. Litan and Singer (2015a, p. 17) report that Vanguard's comparison indicated that 27% of self-directed investors made at least one exchange of money between funds or into other funds and had returns averaging 150 bps lower than those of Target Retirement Funds. In contrast, Litan and Singer indicate the 73% of self-directed investors who did not exchange money (and who by definition did not attempt to time the market) underperformed Target Retirement Funds by 19 bps.⁵⁴ Litan and Singer (1) calculate the weighted average of the underperformance of these two groups relative to Target Retirement Funds ($0.27 \times 150 + 0.73 \times 19 = 54$ bps); (2) assume that the maximum effect of advice that results in avoidance of market timing was this weighed average, while the minimum effect was zero; and (3) pick the mid-point of this assumed range—27 bps—to represent the estimated impact of lost financial advice.

Litan and Singer's use of this estimate as the impact of reduced financial advice is consistent with the *assumptions* that (1) investors currently receiving advice realize returns that approximate the performance of the Vanguard target date funds, i.e., among other things, they do not engage in market timing; (2) reduced financial advice would result in investors who own $27\%/2 = 13.5\%$ of the assets of investors no longer receiving advice making poor timing decisions; and (3) for those investors who previously did not need advice to avoid market timing, reduced financial advice on avoiding market timing would result in investors holding $73\%/2 = 36.5\%$ of assets somehow earning a slightly lower return than they formerly did.

For the portfolio rebalancing estimate, Litan and Singer base their estimates on Vanguard's comparison of the average annual return of a portfolio with 60% stocks and 40% bonds that was not rebalanced over the 53-year period from 1960 to 2013 with the return of a rebalanced portfolio with 80% stocks and 20% bonds. The latter portfolio had about the same risk as the former portfolio, but an average annual return that was 35 bps higher. Litan and Singer's estimate of the effect of reduced advice (1) at least implicitly assumes that Vanguard's comparison of two stylized portfolios is representative of the effect of rebalancing, independent of portfolios that investors actually hold; (2) assumes that the maximum effect of advice that results in better balanced portfolios was the 35 bps spread in the Vanguard comparison, while the minimum effect was zero; and (3) picked the mid-point of this assumed range—17.5 bps—to represent the estimated impact of lost financial advice. Litan and Singer's use of that estimate as the impact of reduced financial advice is consistent with the *assumptions* that (1) investors currently receiving advice realized

⁵⁴ We were unable to find the specific percentages in Vanguard (2014, p. 16). In particular, Vanguard's document indicates that "a majority of investor returns trailed their target-date fund benchmark slightly." While the majority Vanguard describes could be 73%, that specific percentage does not appear in Vanguard (2014).

returns that approximated the performance of Vanguard's balanced portfolio and (2) lost financial advice would result in investors owning 50% of the assets of investors who would lose advice no longer optimally rebalancing.

Losses from Migrating to Advisory Accounts

Litan and Singer assume that some investors will suffer a 31 bps cost increase associated with migrating from brokerage to advisor accounts. This estimate is based on converting Oliver Wyman's (2011, p. 23) single, tersely explained hypothetical example into a number that presumably applies across-the-board to all investors assumed to migrate. In particular, with a minimal amount of detail, Oliver Wyman calculate that a hypothetical 40-year old saver who invested \$25,000 up-front and \$4,000 annually would have 8 percent more savings at age 65 in a brokerage account. That difference, in turn, is equivalent to an annual 31 bps difference.⁵⁵ Therefore, Litan and Singer's *assumption* is that Oliver Wyman's hypothetical example, complete with the excluded costs we described in Section 2 and its static view of the financial industry, provides a reliable estimate of cost increases certain investors could experience.

Evaluation of Litan and Singer's Primary Results

There are several considerations in evaluating the validity of Litan and Singer's primary conclusion—that the 44.5 bps reduction in returns it claims would result from less financial advice and the 31 bps cost from migrating to advisory accounts exceed the 25 basis point gain from less conflicted advice they calculate from DOL's first scenario. These include: (1) the applicable asset base for increases or decreases in annual returns posited by DOL and Litan and Singer; (2) the plausibility of the assumptions Litan and Singer used in applying Vanguard's estimates; (3) whether there are any costs to provide financial advice that are not reflected in Vanguard's examples; and (4) whether Litan and Singer's separately estimated items—market timing and portfolio rebalancing—overlap, i.e., whether the sum of estimates double count some benefits. We discuss these four issues in turn.

With regard to asset base, in order to be informative, Litan and Singer's comparison of their translation of DOL's investor gains and their estimate of annual returns losses from reduced advice would have to address the same asset base. In particular, DOL's analysis related to front-end mutual fund assets, which in principle could include the accounts of investors of various sizes from small investors to much larger investors. In contrast, Litan and Singer appear to limit the potential harms from less financial advice to investors of modest means, e.g., "savers with modest portfolios" (Litan and Singer 2015a, p. 12) and/or "middle-income savers" (Litan and Singer 2015a, p. 16). Litan and Singer provide no information on the asset base of those investors who they believe would lose investment advice if the proposed rule were implemented. While such investors may account for a substantial share of the *accounts*, they account for a much smaller proportion of total assets in front-end load mutual funds. For example, data provided in NERA (2015a, 2015b) suggest that accounts with balances of \$100,000 or less—a threshold that is even higher than the level at which some commenters speculate that investors will lose financial advice—

⁵⁵ $31 \text{ bps} = \sqrt[25]{1 + 0.08} - 1.$

hold about 12% of assets.⁵⁶ In other words, the presumed 44.5 bps impact applies to just 12% of assets. Conversely, the asset base that would apply to the smaller of Litan and Singer's assumed impacts—increased costs from migrating to advisory accounts—would be to the remainder of current brokerage account investors, i.e., on the order of 80% to 90%.⁵⁷

Turning to Litan and Singer's assumptions, for both components they assume (at least implicitly) that (1) the Vanguard examples—Target Retirement Funds in the case of market timing and a stylized balanced portfolio in the case of rebalancing—are representative of the results currently obtained by investors receiving advice and (2) the reduction of advice would result in investors who account for half the asset base making less favorable investment decisions. Litan and Singer do not provide empirical support for either of these assumptions. In addition, because Vanguard (2014a) describes its results as "Vanguard quantifies the value-added of *best practices* in wealth management", even if the Vanguard estimates were representative of the gains from good financial advice, they would be valid only to the extent that all advisers whose services were potentially lost as a result of the Proposed Rule were performing at a best-practices level.⁵⁸ Specifically, Vanguard's advisory services render non-conflicted advice. Litan and Singer attribute the benefits that Vanguard claims for non-conflicted advice to conflicted advice and simply assume away the difference that is the motivation for the proposed regulation.

Further, even if these assumptions were reasonable, the impact for market timing would be overstated because Litan and Singer included the 19 bps difference of self-directed investors who did *not* engage in any market transactions, and by definition could not have been talked out of inadvisable investment, as part of the overall effect. Indeed, Vanguard (2014a) did not appear to view the 19 bps differential as being associated with advice: "The result was that a majority of investor returns trailed their target-date funds slightly, which might be expected based on the funds' expense ratios alone." Seen in this light, rather than being a gain from financial advice, the difference between the 150 bps differential for the investors with transactions and the 19 bps differential for those without transactions, or 131 bps, is the best measure of the impact of market timing implied by Vanguard's results. Correcting Litan and Singer's calculations would (1) reduce the upper bound from 54 bps to 35 bps (0.27×131) and (2) reduce the mid-point of the range from 27 bps to 17.5 bps.

With regard to possible excluded costs associated with financial advice, in describing the benefits from rebalancing Vanguard (2014a, p. 15) notes:

⁵⁶ Similarly, Panis and Brien (2016) show that about 10% of IRA assets are owned by households with IRA assets under \$100,000. ICI (2015a, p. 28) reports a higher percentage—approximately 19%.

⁵⁷ Litan and Singer (2015, pp. 2-4) present a range that is equivalent to a maximum impact of 44.5 bps (everyone losing advice) to a minimum of 31 bps (everyone migrating to brokerage accounts).

⁵⁸ Vanguard's description is consistent with this interpretation: "This paper takes the Advisor's Alpha Framework further by attempting to quantify the benefits that advisors can add *relative to others who are not using such strategies*" (Vanguard 2014, p. 1, emphasis added).

Keep in mind, too, that rebalancing is not necessarily free: There are costs associated with any rebalancing strategy, including taxes and transaction costs, as well as time and labor on the part of advisors. These costs could all potentially reduce your client's return.

Litan and Singer's use of Vanguard's example, without accounting for the additional costs that Vanguard noted, results in an overstatement of possible benefits associated with rebalancing. Further, Vanguard's observation appears to be especially germane to front-end load mutual funds, for which rebalancing would impose relatively high transaction costs.⁵⁹

Finally, with regard to double-counting, the fact that rebalancing is a feature of the Target Retirement Funds Vanguard used to measure the impact of market timing implies that the differential in returns between the target date funds and self-directed investors would capture the effects of both rebalancing and market timing.⁶⁰ Accordingly, as described earlier, to the extent that Vanguard's comparisons are representative of the value of financial advice, the difference between the 150 bps differential between self-directed investors who had market transactions and target retirement funds and the corresponding 19 bps differential for those who had no transactions, or 131 bps, would remove the double-counting with rebalancing. Further, the fact that, as Vanguard explained, the 19 bps differential for the latter group may be explained by expense ratios alone suggests that the benefits from rebalancing may be very small.

In summary, Litan and Singer's bottom line conclusion that the Proposed Rule will lead to a reduction of financial advice that in turn will cost investors more than the benefits DOL estimates would be realized from less conflicted advice is incorrect because (1) Litan and Singer's translation of DOL's investor gains into a basis-point equivalent is incorrectly too low because the estimate improperly mixes real

⁵⁹ Litan and Singer (2015a, p. 17) claim that brokers have an opposite incentive to keep investors in the market so that growing portfolios will produce greater 12b-1 fees. They do not analyze whether the incentive they posit is sufficiently strong to dissuade brokers from advising trades that would produce front-end load shares.

⁶⁰ In their letter to DOL, Litan and Singer (2015b, p. 3) attempt to explain away the double-counting issue as follows:

Mr. Piacentini's fourth criticism is that Vanguard's estimate of the value of portfolio rebalancing reflects some "double-counting" because such rebalancing is already reflected in the performance of the target date funds. In fact, Vanguard's 2014 study makes very clear that its market timing and portfolio rebalancing estimates are different, and the methods used to derive those estimates are also very different.

The fact that Vanguard discussed what it labeled as "best practices" as separate modules and/or used different methods to derive its results says nothing about whether the separate estimates double-count the effects of advice on performance. Litan and Singer's explanation is analogous to a situation in which both rebalancing advice and market timing advice affect returns linearly and a study presents separate simple regression models for each effect. Since some advisors are likely to provide both kinds of advice, the regression coefficient in each model would include the combined effect of both.

estimated gains with a nominal asset base; (2) the asset base to which estimated losses from less financial advice should be applied is a small percentage of the assets held by current brokerage account investors; (3) Litan and Singer incorrectly applied Vanguard's estimates of the value of advice, which related to non-conflicted advice, to conflicted advice; and (4) even if Vanguard's examples of the benefits from financial advice were representative, Litan and Singer's calculations contain errors such as including the differential returns of investors who did not engage in market timing in the estimated effects of market timing and double-counting the effects of market timing and portfolio rebalancing. If Vanguard's examples are informative, the effect of correcting Litan and Singer's errors would be (1) an increase in the gains from DOL's first scenario from 25 bps to 36 bps, (2) elimination of the separate 17.5 bps loss associated with less rebalancing advice, (3) reduction in the loss associated with less advice on avoiding market timing from 27 bps to 17.5 bps, and (4) application of that impact to an asset base no more than 10% to 20% of what DOL calculated, which (5) would result in an impact of about 2 to 3 bps.

Finally, Oliver Wyman's flawed hypothetical example provides no basis for any additional impact associated with putatively higher costs for the remaining investors assumed to migrate to advisory accounts. Also see our discussion of cost differences between brokerage and advisory accounts in Section 2.

Benefits of the Proposed Rule

Litan and Singer briefly criticize DOL's estimate of the benefits from ameliorating conflicted advice, which they routinely characterize as a 25 bps impact. Perhaps as a tacit recognition that their criticisms are rather perfunctory, the authors conclude (Litan and Singer 2015a, p. 23):

[T]he purported 25 basis point gain from the rule claimed by DOL is overstated, most likely by a significant degree. Because the estimated costs of the rule are significantly larger than the purported benefits, there is no need for us to discount the DOL's benefits [...]"

As we now explain, Litan and Singer's criticisms do not undermine the validity of DOL's estimated benefits.

Litan and Singer's criticisms that take issue with DOL's interpretation of such academic studies as Christoffersen et al. (2013) and Bergstresser et al. (2009) overlap those of other parties. For example, NERA (2015a) and ICI (2015a) claim that the age of the data in the academic studies undermines their usefulness, and Berkowitz et al. (2015) claim that the results of Christoffersen et al. (2013) are questionable because their models have low R-squares. The first criticism is misguided, among others because the estimated incentive effects of conflicted compensation are not affected by (declining) average load payments; also see the discussion in Section 2. If anything, ICI (2015c) find that the incentives were sharper in more recent data. The second criticism is similarly misguided, in part because predictors of rates of return are generally elusive and it is therefore noteworthy that conflicted payments had a statistically significant effect on rates of return; see Panis (2015).

Litan and Singer offer three additional criticisms: (1) Christoffersen et al. (2013) suffers from the "fundamental oversight" of estimating underperformance only for the year in which a fund is purchased and of not estimating underperformance "during all the years for which the fund is held," (2) the DOL's overall conclusion on

the negative effects of conflicted advice drawn from Bergstresser et al. (2009) is not consistent with the study's finding that foreign equity funds sold by brokers outperformed foreign equity funds sold through direct channels, and (3) that the "RIA also errs by focusing on the average performance of *funds* rather than of *investors in funds*." None of these criticisms have merit.

First, Litan and Singer's assertion that Christoffersen et al.'s measurement of performance as the forward-looking return for the year following the month in which payments to brokers are observed "does not permit reliable conclusions [...] about any annualized under-performance of funds associated with conflicted advice over the long-run" is at best unhelpful because the authors do not provide an alternative approach that would ameliorate possible problems with Christoffersen et al.'s approach. Absent a well articulated alternative approach, Litan and Singer do not provide the information to determine whether this vague concern has any theoretical, let alone practical, merit. More fundamentally, their concern about long-run performance is misplaced, because the forward-looking return measured at any particular point in time pertains to assets invested not only in the month in question, but also to assets bought (and held) earlier.⁶¹ That is, the return measured by Christoffersen et al. *does* reflect the long-run performance of these earlier-purchased assets.

Second, Litan and Singer's implication that Bergstresser et al.'s finding on the performance of foreign equity funds sold by brokers somehow invalidates DOL's conclusions about conflicted advice overlooks both Bergstresser et al.'s observation that the foreign equity result was hardly typical of their overall findings—"The contrary results in the foreign equity funds are attributable to a single fund complex"—and the authors' subsequent conclusion that "summing up across broad equity, bond, and *foreign equity* investment categories leads us to estimate the annual underperformance of the broker-sold funds at \$4.6 billion in 2004" (Bergstresser et al., 2009, p. 4141, emphasis added).⁶²

Third, while Litan and Singer suggest that a study of investors, rather than funds, could produce results that differ from conclusions drawn from academic studies of fund performance, they provide no empirical evidence to determine whether their concern is of any practical (as opposed to theoretical) importance, or whether the results would become stronger or weaker. In fact, our analysis of information from NERA's investor-level data presented in Section 3 indicates that consistent with the academic studies, the risk-adjusted returns to investors in commission-based accounts lag behind the returns to investors in fee-based accounts.

Litan and Singer's Alternative Disclosure Proposal

Litan and Singer recommend that disclosing the details of how brokers are compensated with a share of a front-end load and ongoing 12b-1 distribution charges is "a more direct and far less costly alternative" to the Proposed Rule. Litan and Singer fault the DOL for relying on a single article by Loewenstein, Cain, and Sah (2011) to conclude that disclosure alone would be insufficient to remedy the harms from conflicted advice. Apparently, they considered that level of support not strong enough to rule out the efficacy of disclosure. Nonetheless, Litan and Singer cite a

⁶¹ The RIA discusses this phenomenon.

⁶² Broker sold funds had a 2004 asset base of \$2.6 trillion, implying a reduction in annual return of 179 basis points (Bergstresser et al., 2009, p. 4136, Table 2).

later article by the same authors (Sah, Loewenstein, and Cain, 2013) as support for the efficacy of their recommended disclosure statement (Litan and Singer 2015a, pp. 26-27):

Yet in subsequent research the very same authors identify conditions under which the burden of disclosure is ameliorated: [...] Three out of four of these conditions would seem to apply to the disclosure remedy proposed above: The disclosure would come from an external source (the Department); the advisee would presumably have the opportunity to change her mind (reinvest her assets) at any point in time; and, the advisee would presumably be able to make the decision in private. Therefore, the disclosure requirements suggested above are consistent with recommendations of the very researchers on which the Department relies.

The factors listed by Sah et al. (2013) appear to *lessen*, but not eliminate, the "burden of disclosure." The burden of disclosure arises when the knowledge that an option that is generally understood to be inferior benefits the adviser at the expense of the advisee actually results in advisees selecting the inferior option more frequently than do advisees who are also advised to select the inferior option, but do not know that the adviser benefits from that choice. For example, while a smaller percentage of advisees selected the inferior alternative recommended by a conflicted adviser when their decision was private, that percentage was still substantially larger than the corresponding advisees to whom the conflict in interest was not disclosed. Further, Sah et al.'s study provides information on both the superior alternative and the inferior alternative that benefits the conflicted adviser; in contrast, Litan and Singer's recommended disclosure statement would describe only the alternative that advantages the conflicted financial adviser. Finally, after reviewing the results of their study as well as related research, Sah et al. (2013, p. 302) conclude:

[T]he optimal solution to COIs [conflicts of interest] is to eliminate them wherever possible, or at least to increase the availability of unbiased advice [...]. The limits of disclosure revealed by these studies and others suggest that policy makers should focus less on disclosing COIs and more on decreasing them.

8. QUANTRIA STRATEGIES

Summary

In response to DOL's Proposed Rule and on behalf of a group of clients, Davis & Harman LLP submitted comments including a study by Quantria Strategies LLC titled "Unintended Consequences: Potential of the DOL Regulations to Reduce Financial Advice and Erode Retirement Readiness" (Quantria 2015). This section contains a review of the Quantria study.

Quantria makes dire predictions about the effects of the Proposed Rule on aggregate retirement outcomes, small businesses, IRA owners, and retirement plan participants. It makes strong assumptions about industry responses. For example, it assumes that financial advisers cannot accept the risk of fiduciary liability and will instead cease to provide advice. The Quantria prediction seems to conflict with the fact that many advisers currently already operate under a fiduciary duty, oftentimes in combination with a lighter standard for some of their other activities.

Quantria assumes that, deprived of financial advice, small businesses will reduce sponsorship of retirement plans and that individuals will increase pension cash-outs, reduce retirement contributions, and commit more investment errors. The benefits that Quantria ascribes to financial advice are based on its persistent confusion of conflicted and non-conflicted advice. While we agree that conflicted advice can confer benefits, overwhelming evidence indicates that they are much smaller than the benefits of non-conflicted advice. Separately, Quantria ignores the possibility that non-conflicted advice could reduce unscrupulous sales practices and root out excessively expensive products. Quantria does not provide any evidence to contradict the possibility that as a result of less conflicted advice, some small business—that currently do not sponsor a retirement plan because of concerns over ulterior motives of their adviser—may start sponsoring a plan.

In short, Quantria relies on unsupported assertions and flawed studies for many of its predictions. Its assumptions about responses by the financial services industry, small businesses, and individuals are unrealistic. As a result, its aggregate estimates of the effects of the Proposed Rule are also unrealistic.

Synopsis

The Quantria study consists of two main parts. The first part discusses unintended effects that Quantria anticipates if the Proposed Rule were implemented as proposed. Quantria argues that the regulation would reduce financial assistance to DC plan participants, owners of small IRAs, and small businesses that may wish to sponsor a retirement plan. Quantria also anticipates a reduction in retirement readiness, i.e., a reduction in such metrics as the fraction of people with sufficient retirement income to cover average expenses and uninsured health care costs (including long-term care costs) at age 65 or older throughout retirement. The authors explain that individuals with lower financial literacy tend to be less prepared for retirement, and that financial advice can help compensate for lack of financial literacy.

The second part discusses anticipated effects of the regulation on retirement savings in more detail for three distinct groups. First, Quantria expects owners of IRAs with low account balances to lose access to financial advice. It also expects fewer IRAs to open as a result of reduced assistance rolling over DC plan balances into an IRA. Second, it expects retirement sponsorship rates among small businesses to fall because of restrictions on marketing activities and because new plans generally carry low balances and may be considered unprofitable to serve. Third, it expects lower DC plan savings because plan participants would have less access to educational materials, would make more investment errors, would take less advantage of employer matching, and would cash-out their DC account upon job separation more often. All combined, Quantria expects losses of retirement savings of \$68 billion to \$80 billion per year. Translated into retirement readiness, "The re-proposed regulations would jeopardize retirement readiness for 11.9 million IRA and retirement participants. This 11.9 million figure consists of individuals who either are unlikely to be retirement ready (6.1 million) or are at risk of failing to be retirement ready (an additional 5.8 million)" (Quantria 2015, p. 32).

Discussion

Overview

The general applicability of the Quantria study hinges on several premises that Quantria adopts. Among these are the following (Quantria 2015, p. 2):

The re-proposed regulations [...] have a general rule that causes many activities of financial advisers to create potential fiduciary liability and they do not provide workable safe harbors in the prohibited transaction exemptions.

And (Quantria 2015, p. 6):

Most importantly, initial indications suggest that very few, if any, financial institutions could satisfy the best interest contract exemption, thereby practically eliminating this exemption.

Quantria offers little or no empirical justification for these assertions. Ultimately, it is an empirical question whether financial institutions will be able to take advantage of prohibited transaction exemptions.

Quantria further asserts the following (Quantria 2015, p. 4):

As a practical matter, financial advisers cannot risk the sanctions imposed if they violate the fiduciary standards, especially the prohibited transaction rules.

Again, Quantria offers little or no empirical justification for this assertion. Many advisers currently operate under a fiduciary standard. In fact, about two-out-of-three advisers already wear two hats, providing financial-planning or portfolio-management services under a fiduciary standard and serving as salespeople of securities, insurance or other products under a lighter duty (Rieker 2015). The assertion is therefore empirically unsupported and inconsistent with current practices.

While the above assertions raise questions, Quantria adopts them as cornerstones of its report. If in fact the forthcoming regulation does provide workable safe harbors in its prohibited transaction exemptions, and if financial advisers in fact are willing to accept fiduciary responsibility, the Quantria study is merely an exercise in hypotheticals.

The remainder of this section parallels the organization of the Quantria study. We first discuss the unintended effects that Quantria anticipates if the Proposed Rule were implemented as proposed (Quantria's Section II). Next we review Quantria's anticipated effects of the regulation on retirement savings (Quantria's Section III).

Unintended Effects of the DOL Regulations

Quantria starts with presenting an overview of the Proposed Rule. With little or no empirical justification, it asserts that the proposed regulations "do not provide workable safe harbors in the prohibited transaction exemptions" (Quantria 2015, p. 2), that "financial advisers cannot risk the sanctions imposed if they violate the fiduciary standards, especially the prohibited transaction rules" (Quantria 2015, p. 4), that "Companies are likely to find that the costs of providing the required information to qualify for the prohibited transaction exemption would exceed the value of getting or retaining a small account" (Quantria 2015, p. 5), and that "initial indications suggest that very few, if any, financial institutions could satisfy the best interest contract exemption" (Quantria 2015, p. 6). Quantria does not specify what these initial indications are or, more generally, what the basis is for its sweeping assertions. At this time, the regulations are not yet in force and it is impossible to tell whether prohibited transaction exemptions will be workable, or whether some advisers will accept fiduciary liability. However, the remainder of Quantria's study hinges on the validity of the above-cited assertions.

Quantria next discusses unintended effects of the Proposed Rule for small businesses, IRA holders, and retirement plan participants.

Small Businesses

Quantria argues that the "DOL regulations likely would reduce the availability of financial assistance for the owners of small account IRAs and small business retirement plans" (Quantria 2015, p. 6). It explains that "financial institutions typically earn different amounts on the different options that a small business can choose to offer its employees. As a result, financial advisers would not be able to provide services to these types of customers" (Quantria 2015, p. 7). In other words, Quantria laments the reduction of conflicted advice. Indeed, that is precisely the purpose of the Proposed Rule. Quantria builds on its premise that financial advisers will not accept fiduciary liability. However, many financial advisers currently operate under a fiduciary standard, and indeed many who avoid fiduciary duty for some of their work accept it for other work (Rieker 2015).

Financial institutions likely have at least two options. They can stop selling retirement plans or they can adapt the compensation of their sales force or intermediaries. Providers who currently rely on front-end load sharing and opaque pricing to sell expensive products may find that they can no longer compete under the new regime. The remaining plan products will likely be less expensive. Under a non-conflicted compensation model, sales people or intermediaries can continue to provide their advisory services. In fact, they would no longer have an incentive to push expensive products, making it easier for them to earn the trust of small

businesses and other clients. A potential outcome of that development is that small businesses—that no longer need to be watchful for unscrupulous sales practices—become more likely to start a retirement plan.

IRA Owners

With respect to existing IRAs, Quantria states (Quantria 2015, p. 7):

Under the regulation, IRA owners would have the following options: (1) if the account is large enough, move to an advisory relationship, which may increase fees, especially for buy and hold investors, (2) if the account is not large enough for an advisory relationship, leave the money in the account, but lose access to an adviser, (3) cash out the savings from the IRA and either spend the money or add the assets to an account that is not tax favored, or (4) roll the IRA assets over to another tax-favored retirement savings account, such as an employer plan, if available.

We discuss these options in turn. First, larger accounts may move to an advisory account. Quantria provides no explanation for its assertion that such accounts may involve higher fees. However, it extensively cites from Oliver Wyman (2011), which tabulated higher fees for fee-based than for commission-based accounts. That comparison accounted for direct costs only and excluded marketing and distribution, shareholder services, and other fees not directly paid by investors. In other words, it excluded load sharing and other indirect payments that are made to commission-based advisers and not to fee-based advisers. Second, Quantria argues that smaller accounts would lose access to an adviser. We discussed and cast serious doubt about this scenario in Section 2. Third, IRA owners could cash out their account, presumably to preserve access to advice. This option is inferior to the other options—particularly in light of recent innovations in the delivery of advice and in such products as target date funds (see Section 2). Fourth, IRA assets may be rolled over into another retirement savings account. The consolidation of assets may indeed be beneficial, as the account holder may be offered lower fees.

In support of its arguments, Quantria cites Garber et al. (2015), which was commissioned by the DOL. For example, “This study, released in February 2015, acknowledges that the regulations could have an adverse effect on some portion of IRA investors and that, as a result, some IRA owners would be worse off under the regulations.” Garber et al. (2015) was indeed released in February 2015, before the Proposed Rule of April 2015. At the time Garber and co-authors wrote their report, the provisions of the Proposed Rule were unknown. Their conclusions stem from general thought experiments of potential effects, not from an analysis of the actual Proposed Rule.

Retirement Plan Participants

Turning to retirement plan participants, Quantria argues that the Proposed Rule would reduce the availability of educational materials for retirement plan participants. It claims that educators would not be allowed to give examples of funds that fit within recommended asset classes. It is our understanding that the Proposed Rule includes a carve-out for educational activities.

Quantria is also concerned that the Proposed Rule would prompt additional cash-outs of DC plans when plan participants terminate their employment. It predicts that their plan’s financial services adviser would stop contacting departing employees because

they would not be allowed to promote their own IRA products. This may not be a bad development, since leaving the assets in the previous employer's plan may be a good option. Current rules require retirement plans to allow terminating employees with balances greater than \$5,000 to stay in the plan (GAO 2014). These employees can continue to enjoy the plan's fiduciary safeguards and access to low-costs funds.

Quantria refers to its 2014 study (Quantria 2014) in which it predicted large increases in cash-outs and large reductions in the lifetime retirement savings as a result of increased cash-outs. Panis (2014) reviewed that study and found it deeply flawed. Among others, Quantria relied on a correlation between financial advice and retirement assets to assert causality, where the causality may in fact go in the reverse direction, and it confused lump sum distributions with cash-outs.

Retirement Readiness

Having discussed unintended consequences for small businesses, IRA holders, and retirement plan participants, Quantria presents the basis for its empirical analysis of unintended effects of the Proposed Rule for retirement outcomes. It offers several definitions of retirement outcomes and adopts the "retirement readiness" measure defined in VanDerHei (2014): "having adequate retirement income to cover average expenses and uninsured health care costs (including long-term care costs) at age 65 or older throughout retirement" (Quantria 2015, p. 9).

Quantria reviews literature on financial literacy and reports that "individuals who lack financial literacy are less likely to plan for retirement and less likely to demonstrate retirement readiness" (p. 11) and that "African-Americans and Hispanics generally have low levels of financial literacy" (Quantria 2015, p. 11).

Quantria argues that low financial literacy can be countered by financial advice. We agree that this is likely the case, but are not convinced by Quantria's arguments. For example:

- Quantria cites Garber et al. (2015) as stating that unsophisticated investors benefit from time savings by using a financial adviser and also from help in choosing investment products. However, Garber et al. (2015) refer to benefits from non-conflicted advisers.
- Quantria also cites Financial Engines/Aon Hewitt (2014), who document that plan participants who benefit from investing in target-date funds, managed accounts and online advice. Again, these are examples of non-conflicted advice.

In other words, Quantria confuses the benefits of non-conflicted advice with those of conflicted advice.

Effects of the Regulations on Overall Retirement Savings

The next part of the Quantria study attempts to quantify the effects of the Proposed Rule on retirement readiness. As before, the authors separately discuss IRA owners, small businesses, and retirement plan participants. In the final stage, Quantria presents aggregate estimates of the Proposed Rule's anticipated effects on retirement readiness.

IRA Owners

Quantria presents general statistics on IRA assets and contributions, distinguishing Traditional, SEP, SIMPLE, and Roth IRAs. The authors point out that “individual or small business investors often open multiple IRA accounts over their retirement savings horizon, creating multiple small accounts” (Quantria 2015, p. 17) that they typically do not consolidate. This suggests that some IRA owners with small balances who may lose access to an adviser will be able to avoid this by consolidating their accounts.

Based on several sources, Quantria estimates such inputs to its model as the fraction of IRAs with small balances and the magnitude of inflows from direct contributions and from rollovers.

Along the way, the authors seek to demonstrate that financial advisers or call centers do not play a major role in encouraging departing employees to roll over their retirement plan assets: “The data on the large number of rollovers exceeds the assistance provided to terminating employees” (sic, Quantria 2015, p. 19). This nonsensical statement appears to be based on rollover activity by 4.1 million taxpayers in 2012 and an unspecified (but presumably smaller) number of contacts with terminating employees.

Quantria relies on Oliver Wyman (2011) for estimates on the number of IRA owners who would lose access to financial advice and the reduction in overall IRA savings. A review by Garber et al. (2015) of the Oliver Wyman study demonstrated that its estimates of account costs (also see above) and industry responses are unreliable. We also reject the predictions of Oliver Wyman (2011) as a reliable basis for any estimates of the consequences of the Proposed Rule.

Quantria also cites Garber et al. (2015) to claim that availability of investment advisers serving the IRA market may well decline after the Proposed Rule takes effect. Indeed, Garber et al. raised that theoretical possibility, but also stated that “[e]ven major reductions in numbers of financial advisors serving the IRA market would not necessarily be economically undesirable, however, because the numbers of professional advisors serving the IRA market currently may be too high from an economic efficiency perspective. Much of the current demand for financial services may be attributable to many retail IRA investors overvaluing these services because these investors do not understand the fees they are paying (directly or indirectly) or the associated costs of advisor self-dealing” (Garber et al., 2015, p. 18). Their argument is similar to the one we made above.

Small Businesses

Quantria presents general statistics on the number of small businesses, the number of people they employ, retirement plan sponsor rates. It relies on a survey by Greenwald and Associates (2014) for estimates of the fraction of small businesses that would stop sponsoring their retirement plan or would reduce employer matches if the Proposed Rule were implemented. The survey describes the Proposed Rule to its respondents in the following language (Greenwald and Associates, 2014, p. 23):

“The Department of Labor is considering prohibiting both retirement plan providers and the advisors who sell retirement plans to employers from assisting the employers in the selection and monitoring of the funds in the retirement plan. Under possible new rules, the employer would have two

options: (a) find an independent expert on investments to provide, for an additional fee, guidance on the selection and monitoring of investment options, or (b) do the selection and monitoring themselves, subject to fiduciary liability if this selection is not done in a prudent manner by someone with sufficient expertise. If "a" is chosen, the plan sponsor would be subject to fiduciary liability if the expert is not chosen in a prudent manner."

This language is patently false and incendiary. The objective of the Proposed Rule is to improve advice, not to reduce it. The Quantria study provides no reliable empirical evidence to the contrary. A potential outcome of non-conflicted advice is that small businesses—that no longer need to be watchful for unscrupulous sales practices—become more likely to sponsor a retirement plan. Although a survey was performed by Greenwald and Associates on this topic, the survey educated respondents in a biased manner. The Greenwald and Associates survey results, therefore, were biased. We reject the survey as a reliable basis for any estimates of the consequences of the Proposed Rule.

Retirement Plan Participants

For current plan participants, Quantria states that about 53% of "[r]etirement plans make available educational information to improve financial literacy as well as investment advice to improve the performance of their employees' retirement plans" (Quantria 2015, p. 25). We agree that such educational information can be beneficial to plan participants. Such assistance is non-conflicted and it is our understanding that the Proposed Rule contains an education carve-out for this purpose.

Quantria also cites Financial Engines/AON Hewitt (2014) and Vanguard (2014b) to argue "that people using managed accounts and online advice have higher average contribution levels than other participants" (Quantria 2015, p. 25). Leaving aside the very likely possibility of reverse causality (people who are serious about retirement are more likely to seek assistance; see Section 2), both studies focus on assistance in the context of DC plans, i.e., this relates to non-conflicted advice. It further cites Financial Engines (2015) as evidence that "participants that use financial advisory services (including both active users of online advice and professional management services) were more likely to maximize their matching contributions" (Quantria 2015, p. 26). Again, this relates to non-conflicted advice. Finally, it argues that DOL (2011) states that "quality advice will address over concentration in employer stock and other failures to properly diversify" (Quantria 2015, p. 27). Again, the DOL (2011) analysis related to (an expansion of) non-conflicted advice.

In short, Quantria repeatedly confuses non-conflicted with conflicted advice. It attributes the benefits of non-conflicted advice to conflicted advice and uses them to estimate the effects of the Proposed Rule. We reject such benefits as a reliable basis for any estimates of the consequences of the Proposed Rule.

For terminating employees, Quantria expects the Proposed Rule to increase retirement plan cash-outs. It confuses gross distributions with cash-outs, as it also did in its earlier study on the subject (Quantria 2014). It states that "42 percent of employees take a cash distribution of their retirement savings at job termination" (Quantria 2015, p. 28) and omits to mention that these cash-outs are overwhelmingly small, accounting for just 7% of dollars (Aon Hewitt 2011). Quantria relies on its 2014 study of cash-outs for estimates of increased cash-outs. As noted above, Panis (2014) reviewed that study and found it deeply flawed. We reject

Quantria (2014) as a reliable basis for any estimates of the consequences of the Proposed Rule.

Retirement Readiness

Based on various assumptions discussed (and rejected) above, Quantria produces projections of the effects of the Proposed Rule on retirement outcomes: "the regulations could be expected to result in losses of retirement savings of \$68-\$80 billion each year" (Quantria 2015, p. 29) and the "regulations would jeopardize retirement readiness for 11.9 million IRA and retirement participants" (Quantria 2015, p. 32). Of these, roughly one-half are labeled "unlikely" to be retirement ready and the other half "at risk of failing" to be retirement ready.

Quantria provided very little detail that would permit an evaluation of its assumptions and calculations. As discussed in detail above, Quantria's assumptions on responses by the financial services industry, small businesses, and individuals are unrealistic. As a result, its aggregate estimates of the effects of the Proposed Rule are also unrealistic.

9. REFERENCES

- Aon Hewitt. 2011. "Leakage of Participants' DC Assets: How Loans, Withdrawals, and Cashouts Are Eroding Retirement Income 2011." Available at http://www.aon.com/attachments/thought-leadership/survey_asset_leakage.pdf.
- AT Kearney. 2015. "Hype vs. Reality: The Coming Waves of `Robo' Adoption." Available at <https://www.atkearney.com/financial-institutions/robo-advisory-services-study>.
- AXA Equitable Life Insurance Company. 2013. "Retirement Cornerstone Variable Annuity Series: Commissions Paid to Broker/Dealer." Available at https://us.axa.com/docs/annuity/rc13_commission_grid_farmers.PDF.
- Bergstresser, Daniel, John Chalmers, and Peter Tufano. 2009. "Assessing the Costs and Benefits of Brokers in the Mutual Fund Industry." *The Review of Financial Studies* 22(10): 4129-4156.
- Berkowitz, Jeremy , Renzo Comolli, and Patrick Conroy. 2015. "Review of the White House Report Titled `The Effects of Conflicted Investment Advice on Retirement Savings'." NERA Economic Consulting, 15 March 2015. http://www.nera.com/content/dam/nera/publications/2015/PUB_WH_Report_Conflicted_Advice_Retirement_Savings_0315.pdf
- Biggs, Andrew G. and Glenn R. Springstead. 2008. "Alternate Measures of Replacement Rates for Social Security Benefits and Retirement Income." *Social Security Bulletin*, Vol. 68(2): 1-19.
- Bluethgen, Ralph, Andreas Gintschel, Andreas Hackethal and Armin Mueller. (2008). "Financial Advice and Individual Investors' Portfolios". Available at SSRN: <http://ssrn.com/abstract=968197>.
- Brady, Peter and Michael Bogdan. 2014. "Who Gets Retirement Plans and Why, 2013." ICI Research Perspective 20, No. 6 (October). Available at www.ici.org/pdf/per20-06.pdf
- Bureau of Labor Statistics. 2014. "National Compensation Survey: Employee Benefits in the United States, March 2014." Available at <http://www.bls.gov/ncs/ebs/benefits/2014/ebbl0055.pdf>.
- Christoffersen, Susan and Richard Evans. 2015. Letter to the Department of Labor, September 21, 2015. Available at <http://www.dol.gov/ebsa/pdf/1210-AB32-2-02766.pdf>.
- Christoffersen, Susan E.K., Richard Evans, and David K. Musto. 2013. "What Do Consumers' Fund Flows Maximize? Evidence from Their Brokers' Incentives." *Journal of Finance* 68(1): 201-235.
- Compass Lexecon. 2015. "Tax Consequences Resulting from Proposed Rules Relating to Financial Representative Fiduciary Status." July 20, 2015. Appendix 5 to the Comments of Primerica. Available at <http://www.dol.gov/ebsa/pdf/1210-AB32-2-00615.pdf>.
- Copeland, Craig. 2014. "Employment-Based Retirement Plan Participation: Geographic Differences and Trends, 2013" Employee Benefit Research Institute Issue Brief 405. Available at http://www.ebri.org/pdf/briefspdf/EBRI_IB_405_Oct14.RetPart.pdf.

- Copeland, Craig. 2015. "Individual Retirement Account Balances, Contributions, and Rollovers, 2013; With Longitudinal Results 2010–2013: The EBRI IRA Database." *EBRI Issue Brief*, no. 414, Employee Benefit Research Institute. Available at http://www.ebri.org/pdf/briefspdf/EBRI_IB_414.May15.IRAs.pdf.
- Council of Economic Advisers. 2015 ("White House report"). "The Effects of Conflicted Investment Advice on Retirement Savings." Available at https://www.whitehouse.gov/sites/default/files/docs/cea_coi_report_final.pdf.
- Del Guercio, Diane, and Jonathan Reuter. 2014. "Mutual Fund Performance and the Incentive to Generate Alpha." *Journal of Finance* 69: 1673-1704.
- DOL. 2011. "Investment Advice—Participants and Beneficiaries." U.S. Department of Labor, *Federal Register*, October 25, 2011, pages 66136-66167.
- DOL. 2015. "Fiduciary Investment Advice: Regulatory Impact Analysis." Available at <http://www.dol.gov/ebsa/pdf/conflictsofinterestria.pdf>.
- Dushi, Irena, Howard M. Iams and Jules Lichtenstein. 2011. "Assessment of Retirement Plan Coverage," *Social Security Bulletin*, Vol. 71(2): 53-65. Available at <https://www.ssa.gov/policy/docs/ssb/v71n2/v71n2p53.pdf>.
- Europe Economics. 2014. "Retail Distribution Review Post Implementation Review." 16 December 2014. Available at <http://www.fca.org.uk/static/documents/research/rdr-post-implementation-review-europe-economics.pdf>.
- Financial Engines. 2015. "Missing out: How Much Employer 401(K) Matching Contributions Do Employees Leave on the Table?" Available at <https://corp.financialengines.com/docs/Financial-Engines-401k-Match-Report-050615.pdf>.
- FutureAdvisor. 2015. "What Is the Account Minimum for the FutureAdvisor Premium Service?", available at <https://futureadvisor.tenderapp.com/kb/going-premium/what-is-the-account-minimum-for-the-futureadvisor-premium-service>.
- GAO. 2014. "401(k) Plans Greater Protections Needed for Forced Transfers and Inactive Accounts." Available at <http://www.gao.gov/assets/670/667151.pdf>.
- Garber, Steven, Jeremy Burke, Angela Hung and Eric Talley. 2015. "Potential Economic Effects on Individual Retirement Account Markets and Investors of DOL's Proposed Rule Concerning the Definition of a 'Fiduciary'." RAND Report RR-1009-DOL, available at <http://www.dol.gov/ebsa/pdf/conflictsofinterestreport6.pdf>.
- Gerhardt, Ralf and Andreas Hackethal. 2009. "The Influence of Financial Advisors on Household portfolios: A study on Private Investors switching to Financial Advice." Goethe-University, Frankfurt am Main.
- Greenwald & Associates. 2014. "The Impact of the Upcoming Re-Proposed Department of Labor Fiduciary Regulation on Small Business Retirement Plan Coverage and Benefits." Available at http://ushcc.com/wp-content/uploads/2014/05/survey_0029436_embargoed_002095743.pdf
- Hedges, Bob, Uday Singh and Monica Gabel. 2015. "Hype vs. Reality: The Coming Waves of "Robo" Adoption." A.T. Kearney.
- ICI. 2006. "401(k) Plans: A 25-Year Retrospective." ICI Research Perspective 12(2).
- ICI. 2015a. Letter from Brian Reid and David W. Blass of the Investment Company Institute to the U.S. Department of Labor, July 21, 2105. Available at <http://www.dol.gov/ebsa/pdf/1210-AB32-2-00749.pdf>.

- ICI. 2015b. Letter from Brian Reid and David W. Blass of the Investment Company Institute to the U.S. Department of Labor, September 24, 2015. Available at <http://www.dol.gov/ebsa/pdf/1210-AB32-2-03056.pdf>.
- ICI. 2015c. Letter from Brian Reid and David W. Blass of the Investment Company Institute to the U.S. Department of Labor, December 1, 2015.
- ICI. 2015d. "2015 Investment Company Fact Book." Investment Company Institute. Available at http://icifactbook.org/pdf/2015_factbook.pdf.
- ICI. 2015e. "The IRA Investor Profile: Traditional IRA Investors' Activity, 2007–2013." Available at https://www.ici.org/pdf/rpt_15_ira_traditional.pdf.
- ICI. 2015f. "The IRA Investor Profile: Roth IRA Investors' Activity, 2007–2013." Available at https://www.ici.org/pdf/rpt_15_ira_roth_investors.pdf.
- International Actuarial Association. 2013. "Stress Testing and Scenario Analysis." Available at http://www.actuaries.org/CTTEES_SOLV/Documents/StressTestingPaper.pdf.
- Investment Funds Institute of Canada. 2012. "The Value of Advice Report 2012." Available at <https://www.ific.ca/wp-content/uploads/2013/02/IFIC-Value-of-Advice-Report-2012.pdf/1650>.
- Jackson National Life Insurance Company. 2013. "Commission Schedule Variable Annuity Contracts." Available at <http://www.cri4.com/crcl/fesco/VAUpdates/Commission Updates/Jackson National VA Commission Update Sept 2013.pdf>.
- Kaplan, Eve. 2012. "9 Reasons You Need To Avoid Variable Annuities," July 2, 2012. Available at <http://www.forbes.com/sites/feeonlyplanner/2012/07/02/9-reasons-you-need-to-avoid-variable-annuities>.
- Kinniry, Francis M., Colleen M. Jaconetti, Michael A. DiJoseph, and Yan Zilbering. 2014. See Vanguard (2014a).
- Kramer, Marc M. 2012. "Financial Advice and Individual Investor Portfolio Performance." *Financial Management* Summer 2012, 395-428.
- Litan, Robert and Hal Singer. 2015a, "Good Intentions Gone Wrong: The Yet-To-Be-Recognized Costs of the Department of Labor's Proposed Fiduciary Rule," July 2015. Available at <http://www.dol.gov/ebsa/pdf/1210-AB32-2-00517.pdf>.
- Litan, Robert and Hal Singer. 2015b. Letter to the U.S. Department of Labor, September 21, 2015. Available at <http://www.dol.gov/ebsa/pdf/1210-AB32-2-02967.pdf>.
- Loewenstein, George, Daylian M. Cain, and Sunita Sah. 2011. "The Limits of Transparency: Pitfalls and Potential of Disclosing Conflicts of Interest." *American Economic Review* 101(3): 423-428.
- Maack, J., 2001. "Scenario Analysis: A Tool for Task Managers." Social Development Paper no. 36. Social Analysis: Selected Tools and Techniques. World Bank, Washington, D.C. Available at <http://siteresources.worldbank.org/EXTSOCIALDEV/Resources/3177394-1167940794463/ScenarioAnalysisMaack.pdf>
- Montmarquette, Claude and Nathalie Viennot-Briot. 2012. "Econometric Models on the Value of Advice of a Financial Advisor." CIRANO. Available at <http://www.cirano.qc.ca/pdf/publication/2012RP-17.pdf>.
- Montmarquette, Claude and Nathalie Viennot-Briot. 2015. "The Value of Financial

- Advice." *Annals of Economics and Finance* 16(1), 69-94.
- NERA. 2015a. "Comment on the Department of Labor Proposal and Regulatory Impact Analysis." NERA Economic Consulting, July 17, 2015. Appendix 1 to the Comments of the Securities Industry and Financial Markets Association, July 20, 2015. Available at <http://www.dol.gov/ebsa/pdf/1210-AB32-2-00506.pdf>.
- NERA. 2015b ("NERA Response"). "Answers to Questions Regarding NERA's 'Comment on the Department of Labor Proposal and Regulatory Impact Analysis.'" NERA Economic Consulting, September 21, 2015. Attached to letter from Kenneth E. Bensten, Jr. to Joseph Piacentini, September 24, 2015. Available at <http://www.dol.gov/ebsa/pdf/1210-AB32-2-03079.pdf>.
- Oliver Wyman, Inc. 2011. "Assessment of the impact of the Department of Labor's proposed 'fiduciary' definition rule on IRA consumers." Available at <http://www.dol.gov/ebsa/pdf/WymanStudy041211.pdf>.
- Oliver Wyman, Inc. 2015. "The Role of Financial Advisors in the U.S. Retirement Market." Available at <http://www.dol.gov/ebsa/pdf/1210-AB32-2-00515.pdf>.
- Panis, Constantijn and Michael J. Brien. 2016. "Financial Asset Holdings of Households in the United States in 2013."
- Panis, Constantijn. 2015. "Comments on a Review of a White House Report on Conflicted Investment Advice." Available at <http://www.dol.gov/ebsa/pdf/conflictsofinterestresearchpaper1.pdf>.
- Panis, Constantijn. 2014. "Review of Study by Quantria Strategies, LLC." Available at <http://www.dol.gov/ebsa/pdf/conflictsofinterestreport2.pdf>.
- Quantria. 2014. "Access to Call Centers and Broker Dealers and Their Effects on Retirement Savings." Quantria Strategies, LLC. Available at http://quantria.com/DistributionStudy_Quantria_4-1-14_final_pm.pdf.
- Quantria. 2015. "Unintended Consequences: Potential of the DOL Regulations to Reduce Financial Advice and Erode Retirement Readiness." July 2015, Attachment to Comments of Davis & Harman, July 21, 2015. Available at <http://www.dol.gov/ebsa/pdf/1210-AB32-2-00746.pdf>.
- Reuter, Jonathan. 2015. "Evaluating the Performance of Broker-Sold Funds." Available at https://www2.bc.edu/jonathan-reuter/research/brokers_revisited_201511.pdf.
- Rieker, Matthias. 2015. "Fiduciary or Broker? Many Financial Advisers Wear Both Hats." *Wall Street Journal*, February 25, 2015.
- Sah, Sunita, George Loewenstein and Daylian M. Cain, 2013. "The Burden of Disclosure: Increased Compliance with Distrusted Advice," *Journal of Personality and Social Psychology* 104(2): 289-304.
- Schrass, Daniel. 2013. "Ownership of Mutual Funds through Investment Professionals, 2012." ICI Research Perspective 19, No. 2 (February). Available at www.ici.org/pdf/per19-02.pdf.
- Scism, Leslie. 2012. "Insurance Fees, Revealed." *Wall Street Journal*, March 30, 2012.
- SEC. 2015. "Mutual Fund Fees and Expenses" webpage of the SEC, available at <http://www.sec.gov/answers/mffees.htm>.
- SigFig. 2016. Available at <https://www.sigfig.com/site/#/pricing>.
- Tergesen, Anne. 2015. "Automated 'Robo' Adviser Lowers Its Investment Minimum."

- Wall Street Journal*, November 20, 2015.
- U.S. Census Bureau. 2015. "Business Dynamics Statistics." Available at http://www.census.gov/ces/dataproducts/bds/data_firm.html.
- UBS. 2015. "Understanding Our Fees, Charges and Other Compensation." Available at https://www.static-ubs.com/us/en/wealth/about_us/fees_compensation/_jcr_content/par/linklist_ee5/link.1510904521.file/bGluay9wYXRoPS9jb250ZW50L2RhbS9zdGF0aWMvd21hbWVyaWNhcy91bmRlcnN0YW5kaW5nX291cl9mZWVzLnBkZg==/understanding_our_fees.pdf.
- VanDerhei, Jack. "What Causes EBRI Retirement Readiness Ratings to Vary: Results from the 2014 Retirement Security Projection Model." Employee Benefit Research Institute, EBRI Issue Brief No. 396, February 2014.
- Vanguard. 2014a. "Putting a Value on Your Value: Quantifying Vanguard Advisor's Alpha," March 2014. Available at <http://www.vanguard.com/pdf/ISGQVAA.pdf>.
- Vanguard. 2014b. "How America Saves 2014: A Report on Vanguard 2013 Defined Contribution Plan Data." Available at <https://institutional.vanguard.com/iam/pdf/HAS14.pdf>.
- White House Report. 2015. See Council of Economic Advisers (2015).

DISCLAIMER

The views, opinions, and/or findings contained in this report should not be construed as an official Government position, policy or decision, unless so designated by other documentation issued by the appropriate governmental authority.

We call your attention to the possibility that other professionals may perform procedures concerning the same information or data and reach different findings than Advanced Analytical Consulting Group, Inc. (AACG) for a variety of reasons, including the possibilities that additional or different information or data might be provided to them that was not provided to AACG, that they might perform different procedures than did AACG, or that professional judgments concerning complex, unusual, or poorly documented matters may differ.

This document contains general information only. AACG is not, by means of this document, rendering business, financial, investment, or other professional advice or services. This document is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action. Before making any decision or taking any action, a qualified professional adviser should be consulted. AACG, its affiliates, or related entities shall not be responsible for any loss sustained by any person who relies on this publication.

This document represents the final versions of the third and fourth Critical Reviews pursuant to Task Order DOL-OPS-15-T-00059 and the first, second, third, and fourth Critical Reviews pursuant to Task Order DOL-OPS-15-T-00186 under Contract DOL-OPS-14-D-0021.

79

Rates of Return of Broker-Sold and Direct-Sold Mutual Funds

March 15, 2016

Advanced Analytical Consulting Group, Inc.

Karthik Padmanabhan, MBA, MS
312-551-9001
karpad@aacg.com

Constantijn Panis, PhD
213-784-6400
stanpanis@aacg.com

Timothy J. Tardiff, PhD
617-827-4043
timtardiff@aacg.com

ABSTRACT

The U.S. Department of Labor (DOL) is analyzing historical mutual fund returns to measure the performance of funds sold directly to investors compared to funds sold via a broker-dealer. Using data from Morningstar, a widely used source of research on mutual funds, the DOL is measuring yearly performance from 1980 to 2015 for two asset types – domestic equities and international equities.

Advanced Analytical Consulting Group (AACG) has independently replicated the results obtained by the DOL. In this report we discuss the effort, the data acquisition, analysis, results and potential data issues that may affect the findings.

CONTENTS

ABSTRACT	i
1. INTRODUCTION.....	1
2. DESCRIPTION	1
3. DATA ANALYSIS	1
Data Acquisition.....	1
Data Analysis	8
Results	9
4. POTENTIAL DATA ISSUES.....	13
Incomplete Data in Morningstar	13
Using Share Class Type to Determine Sales Channel	13
APPENDIX. Share Class Type Descriptions	14
Share Classes Associated with Broker-Sold Funds	14
Share Classes Associated with Direct-Sold Funds	15
Share Classes Associated with Neither Broker- Nor Direct-Sold Funds	15
Disclaimer	17

1. INTRODUCTION

The U.S. Department of Labor (DOL) is performing analysis on Morningstar data to measure relative performance between two groups of mutual funds. Advanced Analytical Consulting Group (AACG) provided assistance by independently replicating the results obtained by DOL and highlighting potential data issues that influence the results. This document discusses AACG's work on this task.

2. DESCRIPTION

The DOL is measuring the yearly performance from 1980 to 2015 of mutual funds sold by broker-dealers relative to mutual funds that are directly sold, for two asset types—domestic equities (including sector funds) and international equities. DOL obtained Morningstar data for this effort through the Morningstar Direct product offering. Morningstar is one of the best known sources for research on mutual funds.

3. DATA ANALYSIS

Data Acquisition

The data used by AACG were retrieved using the Morningstar Direct product offering. Morningstar Direct allows a user to retrieve specific data elements for a group of mutual funds and export the results to Microsoft Excel.

The following search criteria were used to retrieve data for the two asset types. The checkbox for "Only Surviving Investment" was left unchecked to include data for funds that are no longer active.

Figure 1. Domestic Equity Open-End Funds (1 of 3)

Morningstar Direct -- Open End Funds

Open Insert Delete Clear All Export PDF Save As Feedback

	Rel	(Field Name	Operator	Value)
<input type="checkbox"/>	1	---	Domicile	=	United States	
<input type="checkbox"/>	2	And	(
<input type="checkbox"/>	3	Or	Global Category	=	US Equity Large Cap Blend	
<input type="checkbox"/>	4	Or	Global Category	=	US Equity Large Cap Growth	
<input type="checkbox"/>	5	Or	Global Category	=	US Equity Large Cap Value	
<input type="checkbox"/>	6	Or	Global Category	=	US Equity Mid Cap	
<input type="checkbox"/>	7	Or	Global Category	=	US Equity Mid/Small Cap	
<input type="checkbox"/>	8)
<input type="checkbox"/>	9					
<input type="checkbox"/>	10					

View Field Name: By Category Alphabetically

Include: Only Surviving Investments User Defined Primary Class Only

Items Searched: 0 Items Found: -

Run Search

Help OK Cancel

Figure 2. Domestic Equity Open-End Funds (2 of 3)

Morningstar Direct -- Open End Funds

Open Insert Delete Clear All Export PDF Save As Feedback

	Rel	(Field Name	Operator	Value)
<input type="checkbox"/>	1	---	Domicile	=	United States	
<input type="checkbox"/>	2	And	(
<input type="checkbox"/>	3	Or	Global Category	=	Communications Sector Equity	
<input type="checkbox"/>	4	Or	Global Category	=	Consumer Goods & Services Sector	
<input type="checkbox"/>	5	Or	Global Category	=	Energy Sector Equity	
<input type="checkbox"/>	6	Or	Global Category	=	Financials Sector Equity	
<input type="checkbox"/>	7	Or	Global Category	=	Healthcare Sector Equity	
<input type="checkbox"/>	8	Or	Global Category	=	Industrials Sector Equity	
<input type="checkbox"/>	9	Or	Global Category	=	Natural Resources Sector Equity	
<input type="checkbox"/>	10	Or	Global Category	=	Other Sector Equity	
<input type="checkbox"/>)

View Field Name: By Category Alphabetically

Include: Only Surviving Investments User Defined Primary Class Only

Items Searched: 0 Items Found: -

Run Search

Help OK Cancel

Figure 3. Domestic Equity Open-End Funds (3 of 3)

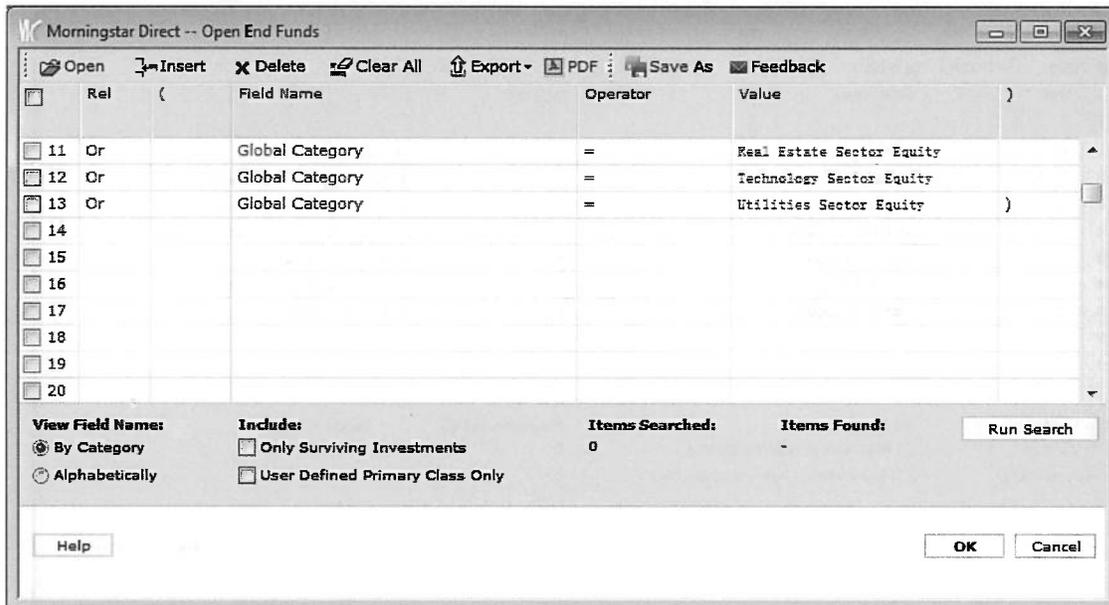
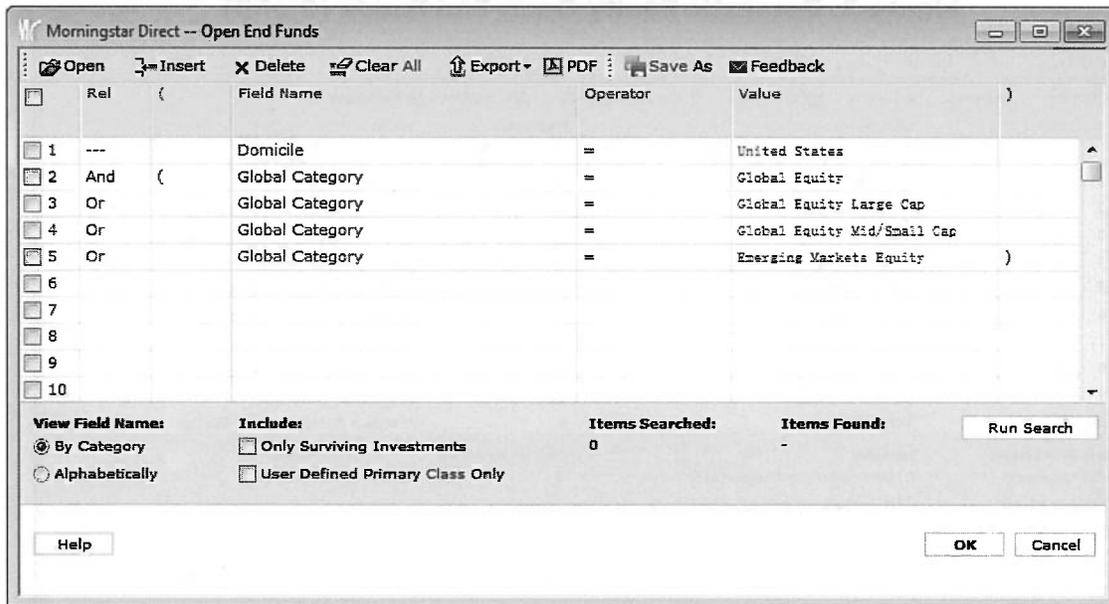


Figure 4. International Equity Open-End Funds



The following 16 data fields were retrieved from each search query.

Field Name	Morningstar Description and Calculation Method
Name	The name of the investment.
Ticker	The identifier under which a security trades on an exchange.

Field Name	Morningstar Description and Calculation Method
Morningstar Category	In an effort to distinguish funds by what they own, as well as by their prospectus objectives and styles, Morningstar developed the Morningstar Categories. While the prospectus objective identifies a fund's investment goals based on the wording in the fund prospectus, the Morningstar Category identifies funds based on their actual investment styles as measured by their underlying portfolio holdings (portfolio and other statistics over the past three years). See specific category name for further details (e.g., "Category - Large Value").
Inception Date	Date on which the security is first offered.
Net Assets Date	The as of date for the fund share class' net assets.
Net Assets – Share Class	The net assets of the mutual fund, recorded in unit of base currency. Net-asset figures are useful in gauging a fund's size, agility, and popularity. They help determine whether a small company fund, for example, can remain in its investment-objective category if its asset base reaches an ungainly size.
Fund Size Date	The as of date of fund size (asset under management) of a fund.
Fund Size	The total amount of money managed as a standalone portfolio across share classes/subaccounts. Fund Size is useful in gauging a product's size, agility, and popularity. This can be greater than or equal to the share class/subaccount net assets. (They will be equal if only one share class is offered or the fund only appears in one policy).
Annual Report Net Expense Ratio	The percentage of fund assets used to pay for operating expenses and management fees, including 12b-1 fees, administrative fees, and all other asset-based costs incurred by the fund, except brokerage costs. Fund expenses are reflected in the fund's NAV. Sales charges are not included in the expense ratio. The expense ratio for fund of funds only includes the wrap or sponsor fees, and does not include the underlying fund fees.
Annual Report Gross Expense Ratio	The Gross Expense Ratio represents the total gross expenses (net expenses with waivers added back in) divided by the fund's average net assets. If it is not equal to the net expense ratio, the gross expense ratio portrays the fund's expenses had the fund not waived a portion, or all, of its fees. Thus, to some degree, it is an indication of fee contracts. Some fee waivers have an expiration date; other waivers are in place indefinitely.
12b-1 Fee	The maximum annual charge deducted from fund assets to pay for distribution and marketing costs. Although usually set on a percentage basis, this amount will occasionally be a flat figure. Only active 12b-1 plans are represented here. This information is

Field Name	Morningstar Description and Calculation Method
	taken directly from the fund's prospectus. (Morningstar lists the maximum amount.)
Share Class Type	Indicates the share class for open-end funds. Shares of the same fund that offer different shareholder rights and obligations, such as different fee and load charges. Common share classes are A (front-end load), B (deferred fees), C (no sales charge and a relatively high annual 12b-1 fee, such as 1.00%). Multi-class funds hold the same investment portfolio for all classes, and differ only in their surrounding fee structure.
Return by Month (1980 to 2015) Return by Year (1980 to 2015)	Expressed in percentage terms, Morningstar's calculation of total return is determined each month by taking the change in monthly net asset value, reinvesting all income and capital-gains distributions during that month, and dividing by the starting NAV. Reinvestments are made using the actual reinvestment NAV, and daily payoffs are reinvested monthly. Unless otherwise noted, Morningstar does not adjust total returns for sales charges (such as front-end loads, deferred loads and redemption fees), preferring to give a clearer picture of a fund's performance. The total returns do account for management, administrative, 12b-1 fees and other costs taken out of fund assets. Total returns for periods longer than one year are expressed in terms of compounded average annual returns (also known as geometric total returns), affording a more meaningful picture of fund performance than non-annualized figures.
Net Assets – Share Class by Month (1980-01 to 2015-12)	Monthly share-class level total net assets.
Alpha by Year (1980 to 2015)	<p>A measure of the difference between a portfolio's actual returns and its expected performance, given its level of risk as measured by beta. A positive Alpha figure indicates the portfolio has performed better than its beta would predict. In contrast, a negative Alpha indicates the portfolio has underperformed, given the expectations established by beta.</p> <p>Alpha is calculated by taking the excess average monthly return of the investment over the risk free rate and subtracting beta times the excess average monthly return of the benchmark over the risk free rate. The equation is as follows:</p> $\alpha_M = \bar{R}^e - \beta \bar{B}^e$ <p>where</p> <p>α_M = Monthly measure of alpha</p> <p>\bar{R}^e = Average monthly excess return of the investment</p>

Field Name	Morningstar Description and Calculation Method
	\bar{B}^e = Average monthly excess return of the benchmark β = Beta <p>The resulting alpha is in monthly terms, because the average returns for the portfolio and benchmark were monthly averages. Morningstar then multiplies it by 12 to put it in annual terms.</p> $\alpha_A = 12 \alpha_M$ <p>The same method applies for alpha (non-excess return) except that the raw return is used instead of the excess return.</p>

By executing the search queries described above, AACG obtained records for the following numbers of funds.¹

	<i>Domestic Equity Funds</i>	<i>International Equity Funds</i>
Date extracted	Feb 5, 2016	Feb 5, 2016
Number of funds retrieved	22,316	7,160

In addition to the alpha performance metric computed and published by Morningstar, we calculated two additional performance metrics – the one-factor alpha and the three-factor alpha for each fund and each calendar year with complete data (12 monthly returns). Like the Morningstar alpha, both metrics measure the difference between the fund's return and its expected performance given its level of risk.

We estimated one-factor alphas through a regression analysis of excess market returns on excess fund returns. Similarly, we estimated three-factor alphas through a regression analysis of three independent variables—the excess market returns, the Small minus Big (SMB) factor and the High minus Low (HML) factor. We used market returns, risk-free rates, SMB factors, and HML factors developed by Eugene Fama and Kenneth French.² The SMB factor is the return of a portfolio of small stocks in excess of the return of a portfolio of large stocks. The HML factor is the return of a portfolio of stocks with high ratios of book value to market value in excess of the return on a portfolio of stocks with low book-to-market ratios.³

¹ Domestic equity funds as extracted from Morningstar Direct included approximately 10 funds with "EAA" investment categories, which suggested their investment focus and portfolio holdings are concentrated in Europe, Asia and Africa. We reclassified these funds as International.

² http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html. Factors for Domestic Equity funds were obtained from "U.S. Research Returns Data" and Factors for International Equity Funds were obtained from "Developed Market Factors and Returns: Fama/French Global ex US Factors."

³ Zvi Bodie, Alex Kane, and Alan J. Marcus. 2001. *Investments*. 5th edition, McGraw-Hill Education.

The resulting estimate of alpha is in monthly terms, because the returns for the fund and benchmark were monthly returns. An annual alpha was calculated using the following equation.⁴

$$\alpha_A = (1 + r_m)^{12} - (1 + r_m - \alpha_m)^{12}$$

where

r_m = Average monthly fund return

α_m = Monthly measure of alpha

⁴ For intuition: α_m is the average monthly amount by which a fund's risk-adjusted return exceeds the benchmark return, and $r_m - \alpha_m$ may thus be interpreted as the benchmark average monthly return. Both the fund's average monthly return and the benchmark average monthly return are compounded into annual values; their difference represents a fund's annual alpha.

Data Analysis

For each asset type and year, we aggregated estimated alphas by distribution channel, weighted by fund assets.

- i. For each fund, an average asset size for each year was determined by calculating a simple average of 12 monthly values of "Net Assets – Share Class by Month".⁵
- ii. Using the "Share Class Type" field and convention described below, each fund was classified as sold by a broker-dealer, directly, or neither. ("Neither" is not used in subsequent analysis.)

Share Class Type	Classification
A Adv B C D T	Broker-Dealer Sold
Inv No Load S	Direct-Sold
(BLANK) Inst Load Waived M N Other Retirement	Neither

For a detailed description of the share class type, see the Appendix.

- iii. For pooled broker-dealer funds and pooled direct-sold funds, a weighted average annual return and weighted-average alpha was calculated by calendar year. The weighting used the average asset size described above.
- iv. A small number of funds did not have asset size data for some or all years. An average asset size for the year could not be calculated. These funds were removed from the weighted average calculations. (Also see below.)
- v. Some funds did not have return and/or alpha information for some or all years. In such cases, these funds were not included in the calculation of the weighted average returns and alphas respectively.

⁵ At the beginning and end of a fund's life cycle, fewer than 12 monthly asset values were typically available. For such partial years, the Morningstar data did not contain information on alpha. We similarly excluded partial years from our analysis.

Table 1 shows the number of funds that were used to calculate average returns and average alphas.

Table 1. Number of Funds Used in the Analysis

	Domestic		International	
	Raw	Used	Raw	Used
Broker-sold	8,597	7,150	2,661	2,189
Direct-sold	1,880	1,710	456	409
Neither	11,839		4,043	
Total	22,316	8,860	7,160	2,598

Results

Tables 2, 3, and 4 show average returns and average risk-adjusted returns for domestic funds, international funds, and pooled domestic/international funds.

For each measure of rates of (risk-adjusted) returns, the tables list asset-weighted average returns for broker-sold funds, for direct-sold funds, and their difference ("overperformance"). This difference captures the extent to which broker-sold funds outperformed direct-sold funds. Negative numbers indicate underperformance of broker-sold funds.

Table 2. Average Rates of Return and Alphas of Domestic Equity Open-End Funds

Year	Raw Returns			Morningstar's Alpha			One-Factor Alpha			Three-Factor Alpha		
	Broker-sold	Direct-sold	Overperformance	Broker-sold	Direct-sold	Overperformance	Broker-sold	Direct-sold	Overperformance	Broker-sold	Direct-sold	Overperformance
1980	33.75%	33.93%	-0.18%	-0.31%	-0.62%	0.32%	1.54%	1.11%	0.42%	-0.30%	-0.48%	0.18%
1981	-3.24%	-2.30%	-0.94%	3.17%	5.02%	-1.85%	-0.21%	1.57%	-1.78%	0.31%	1.08%	-0.77%
1982	27.01%	21.45%	5.56%	4.67%	0.12%	4.55%	5.93%	0.94%	4.99%	5.01%	0.29%	4.72%
1983	20.07%	24.65%	-4.58%	-0.29%	1.59%	-1.87%	-1.40%	1.40%	-2.80%	-0.16%	1.88%	-2.05%
1984	-0.94%	-0.58%	-0.36%	-6.48%	-5.77%	-0.70%	-4.06%	-3.25%	-0.81%	-1.90%	-1.95%	0.05%
1985	27.44%	29.08%	-1.64%	-3.75%	-1.35%	-2.41%	-3.85%	-1.87%	-1.99%	-2.65%	-1.35%	-1.29%
1986	17.11%	17.16%	-0.05%	-0.63%	-0.03%	-0.61%	2.11%	2.08%	0.03%	1.36%	2.56%	-1.20%
1987	3.57%	3.31%	0.26%	-2.66%	-2.30%	-0.36%	1.81%	1.99%	-0.18%	2.76%	3.59%	-0.83%
1988	14.04%	18.18%	-4.14%	-0.65%	2.75%	-3.41%	-2.16%	1.08%	-3.25%	0.25%	1.12%	-0.87%
1989	27.30%	26.65%	0.65%	-1.37%	-1.72%	0.35%	0.27%	-0.85%	1.12%	0.62%	0.22%	0.40%
1990	-4.85%	-6.72%	1.86%	-2.01%	-3.84%	1.83%	-0.03%	-0.95%	0.92%	0.89%	1.30%	-0.41%
1991	31.92%	38.27%	-6.35%	2.35%	4.83%	-2.48%	-0.39%	2.15%	-2.55%	-1.95%	-0.89%	0.31%
1992	7.51%	8.65%	-1.14%	0.69%	1.77%	-1.09%	-1.74%	-0.73%	-1.01%	-1.95%	-2.38%	0.44%
1993	13.94%	17.65%	-3.71%	3.45%	6.38%	-2.93%	3.54%	6.81%	-3.27%	2.25%	4.14%	-1.89%
1994	-1.46%	-0.32%	-1.13%	-2.66%	-1.69%	-0.97%	-1.20%	0.06%	-1.26%	-1.04%	0.36%	-1.40%
1995	32.78%	33.48%	-0.69%	-1.97%	-1.78%	-0.18%	-4.76%	-7.63%	2.87%	-0.63%	0.99%	-1.61%
1996	18.05%	19.04%	-0.98%	-2.55%	-1.77%	-0.77%	-2.30%	-0.57%	-1.73%	-1.39%	-1.06%	-0.33%
1997	25.37%	26.17%	-0.80%	-3.12%	-2.31%	-0.80%	-4.00%	-3.34%	-0.66%	-5.66%	-5.38%	-0.28%
1998	20.65%	21.86%	-1.21%	-7.27%	-6.33%	-0.94%	-2.80%	-1.57%	-1.23%	-1.61%	-2.67%	1.06%
1999	29.68%	24.90%	4.78%	4.75%	1.96%	2.80%	5.44%	1.94%	3.50%	1.81%	1.30%	0.51%
2000	-7.14%	-4.59%	-2.55%	3.26%	6.18%	-2.92%	6.46%	7.67%	-1.21%	5.60%	7.11%	-1.50%
2001	-14.32%	-10.62%	-3.71%	-0.32%	2.89%	-3.21%	-1.81%	0.88%	-2.69%	-3.04%	-0.82%	-2.22%
2002	-22.30%	-20.10%	-2.20%	-2.95%	-0.03%	-2.91%	-2.06%	0.22%	-2.28%	-2.68%	-0.37%	-2.31%
2003	30.80%	31.44%	-0.64%	1.65%	2.54%	-0.88%	-0.62%	0.64%	-1.25%	-2.81%	-1.43%	-1.38%
2004	11.73%	12.74%	-1.01%	-0.83%	0.09%	-0.92%	-0.67%	0.38%	-1.06%	-0.22%	0.88%	-1.09%
2005	7.94%	8.18%	-0.24%	2.42%	2.46%	-0.04%	1.90%	2.10%	-0.20%	1.20%	1.35%	-0.15%
2006	13.67%	14.08%	-0.41%	-4.07%	-4.24%	0.17%	-2.75%	-2.88%	0.13%	-4.32%	-3.79%	-0.54%
2007	7.62%	8.13%	-0.51%	1.38%	1.70%	-0.33%	1.97%	2.46%	-0.50%	1.45%	1.88%	-0.43%
2008	-38.88%	-39.35%	0.47%	1.99%	4.15%	-2.16%	-0.14%	1.14%	-1.28%	0.32%	0.98%	-0.67%
2009	32.60%	34.22%	-1.62%	5.81%	5.89%	-0.07%	6.00%	6.01%	-0.01%	4.81%	5.26%	-0.45%
2010	16.20%	18.33%	-2.13%	1.24%	2.75%	-1.51%	-3.08%	0.88%	-1.55%	-0.73%	-0.49%	-0.24%
2011	-2.88%	-1.77%	-1.11%	-4.17%	-3.05%	-1.13%	-3.07%	-1.96%	-1.11%	-3.42%	-2.34%	-1.07%
2012	14.99%	16.13%	-1.13%	-0.47%	-0.47%	-0.47%	-1.30%	-0.84%	-0.46%	-1.50%	-0.66%	-0.84%
2013	32.11%	33.89%	-1.78%	1.72%	2.35%	-0.63%	-0.28%	0.03%	-0.31%	-3.36%	-2.39%	-0.98%
2014	9.75%	10.72%	-0.97%	-3.23%	-2.31%	-0.92%	-1.64%	-0.76%	-0.88%	-1.06%	0.45%	-1.51%
2015	-0.54%	0.59%	-1.13%	-1.64%	-0.43%	-1.21%	-0.49%	0.65%	-1.15%	-1.33%	-0.58%	-0.75%

Source: AACG calculations based on Morningstar data.

Table 3. Average Rates of Return and Alphas of International Equity Open-End Funds

Year	Raw Returns			Morningstar's Alpha			One-Factor Alpha			Three-Factor Alpha		
	Broker-sold	Direct-sold	Overperformance	Broker-sold	Direct-sold	Overperformance	Broker-sold	Direct-sold	Overperformance	Broker-sold	Direct-sold	Overperformance
1980	28.75%	23.72%	5.03%									
1981	1.16%	-3.10%	4.26%									
1982	18.08%	8.29%	9.78%									
1983	31.94%	31.64%	0.30%									
1984	0.78%	-0.66%	1.44%									
1985	30.30%	44.36%	-14.07%									
1986	22.74%	59.25%	-36.51%									
1987	5.11%	11.59%	-6.49%									
1988	19.09%	14.19%	4.91%									
1989	23.12%	23.44%	-0.32%									
1990	-9.88%	-8.90%	-0.98%									
1991	26.96%	11.94%	15.02%									
1992	3.03%	-3.49%	6.51%									
1993	30.41%	38.97%	-8.56%									
1994	-1.06%	-1.79%	0.73%									
1995	17.08%	11.71%	5.37%									
1996	17.75%	16.56%	1.19%									
1997	11.57%	7.96%	3.61%									
1998	8.53%	15.62%	-7.09%									
1999	39.74%	47.12%	-7.39%									
2000	-9.72%	-14.47%	4.76%									
2001	-11.84%	-19.42%	7.57%									
2002	-14.54%	-17.60%	3.06%									
2003	35.72%	36.39%	-0.67%									
2004	17.92%	18.09%	-0.18%									
2005	14.43%	17.14%	-2.70%									
2006	22.41%	25.40%	-2.99%									
2007	14.65%	15.76%	-1.11%									
2008	-41.97%	-45.76%	3.79%									
2009	38.93%	40.63%	-1.69%									
2010	12.57%	12.98%	-0.41%									
2011	-10.84%	-14.25%	3.41%									
2012	18.12%	19.84%	-1.72%									
2013	22.18%	20.66%	1.52%									
2014	-0.07%	-2.13%	2.05%									
2015	-2.08%	-4.05%	1.97%									

The international benchmark returns from Fama-French start in July 1990

Source: AACG calculations based on Morningstar data.

Table 4. Average Rates of Return and Alphas of Domestic and International Equity Open-End Funds

Year	Raw Returns			Morningstar's Alpha			One-Factor Alpha			Three-Factor Alpha		
	Broker-sold	Direct-sold	Overperformance	Broker-sold	Direct-sold	Overperformance	Broker-sold	Direct-sold	Overperformance	Broker-sold	Direct-sold	Overperformance
1980	33.14%	33.85%	-0.71%	-0.35%	-0.62%	0.27%	1.54%	1.11%	0.42%	-0.30%	-0.48%	0.18%
1981	-2.67%	-2.32%	-0.35%	3.05%	5.02%	-1.97%	-0.21%	1.57%	-1.78%	0.31%	1.08%	-0.77%
1982	25.68%	21.18%	4.50%	4.77%	0.12%	4.65%	5.93%	0.94%	4.99%	5.01%	0.29%	4.72%
1983	21.92%	24.78%	-2.86%	0.32%	1.59%	-1.26%	-1.40%	1.40%	-2.80%	-0.16%	1.88%	-2.05%
1984	-0.65%	-0.58%	-0.07%	-6.15%	-5.77%	-0.38%	-4.06%	-3.25%	-0.81%	-1.90%	-1.95%	0.05%
1985	27.97%	29.78%	-1.81%	-4.42%	-1.35%	-3.08%	-3.85%	-1.87%	-1.99%	-2.65%	-1.35%	-1.29%
1986	18.12%	21.08%	-2.96%	-1.05%	-0.03%	-1.02%	2.11%	2.08%	0.03%	1.36%	2.56%	-1.20%
1987	3.85%	4.09%	-0.25%	-3.47%	-2.30%	-1.17%	1.81%	1.99%	-0.18%	2.76%	3.59%	-0.83%
1988	15.02%	17.94%	-2.92%	-0.03%	2.75%	-2.79%	-2.16%	1.08%	-3.25%	0.25%	1.12%	-0.87%
1989	26.41%	26.48%	-0.06%	-0.64%	-1.72%	1.08%	0.27%	-0.85%	1.12%	0.62%	0.22%	0.40%
1990	-5.95%	-6.86%	0.91%	-2.10%	-3.84%	1.73%	-0.03%	-0.95%	0.92%	0.89%	1.30%	-0.41%
1991	30.86%	36.72%	-5.85%	3.23%	4.83%	-1.60%	3.56%	2.21%	1.35%	3.18%	-0.67%	3.85%
1992	6.61%	8.02%	-1.41%	0.90%	1.77%	-0.88%	-0.63%	-0.59%	-0.04%	-0.27%	-2.00%	1.73%
1993	17.23%	19.08%	-1.85%	3.90%	6.38%	-2.48%	6.53%	7.81%	-1.28%	5.47%	5.00%	0.47%
1994	-1.36%	-0.48%	-0.88%	-2.83%	-1.69%	-1.14%	-2.89%	-0.92%	-1.96%	-0.72%	0.28%	-1.00%
1995	28.95%	31.61%	-2.66%	-1.58%	-1.78%	0.21%	-1.21%	-6.65%	5.43%	1.85%	1.24%	0.62%
1996	17.98%	18.83%	-0.85%	-1.97%	-1.77%	-0.20%	1.01%	0.34%	0.67%	1.33%	-0.28%	1.60%
1997	22.10%	24.57%	-2.47%	-2.94%	-2.31%	-0.63%	-0.48%	-2.32%	1.84%	-2.89%	-4.85%	1.96%
1998	17.96%	21.34%	-3.38%	-7.66%	-6.33%	-1.33%	-3.79%	-1.51%	-2.28%	-2.12%	-2.39%	0.27%
1999	31.61%	26.69%	4.92%	5.22%	2.52%	2.70%	3.83%	1.73%	2.10%	-0.71%	0.55%	-1.27%
2000	-7.62%	-5.54%	-2.08%	3.61%	5.99%	-2.38%	6.63%	7.54%	-0.91%	6.31%	7.64%	-1.33%
2001	-13.87%	-11.39%	-2.47%	1.28%	2.82%	-1.54%	0.08%	1.09%	-1.01%	-0.73%	-0.32%	-0.41%
2002	-20.85%	-19.88%	-0.96%	-2.44%	-0.35%	-2.08%	-2.13%	-0.24%	-1.89%	-1.02%	-0.19%	-0.83%
2003	31.75%	31.87%	-0.12%	0.97%	1.96%	-0.99%	-2.03%	-0.38%	-1.65%	-4.80%	-2.48%	-2.32%
2004	13.04%	13.25%	-0.21%	-1.00%	-0.17%	-0.83%	-1.29%	-0.14%	-1.15%	-1.20%	0.14%	-1.35%
2005	9.49%	9.22%	0.27%	1.94%	2.30%	-0.35%	1.21%	1.81%	-0.60%	1.06%	1.26%	-0.20%
2006	16.06%	15.84%	0.22%	-3.51%	-3.88%	0.37%	-2.55%	-2.86%	0.31%	-4.70%	-4.03%	-0.67%
2007	9.73%	9.58%	0.15%	0.91%	1.24%	-0.33%	2.03%	2.51%	-0.48%	1.10%	1.94%	-0.83%
2008	-39.83%	-40.64%	0.81%	0.91%	3.18%	-2.27%	-0.86%	0.56%	-1.42%	-0.78%	0.32%	-1.10%
2009	34.50%	35.48%	-0.97%	4.90%	4.39%	0.51%	7.20%	6.06%	1.14%	7.92%	6.27%	1.65%
2010	15.09%	17.26%	-2.17%	1.37%	2.47%	-1.10%	0.09%	0.96%	-0.87%	0.24%	0.20%	0.04%
2011	-5.27%	-4.27%	-1.00%	-2.57%	-2.51%	-0.05%	-1.75%	-1.72%	-0.02%	-3.19%	-2.74%	-0.45%
2012	15.88%	16.83%	-0.95%	0.33%	0.25%	0.08%	0.01%	-0.15%	0.16%	0.07%	0.07%	0.00%
2013	29.32%	31.21%	-1.89%	3.21%	2.97%	0.23%	0.93%	-0.03%	0.97%	-0.86%	-1.69%	0.83%
2014	7.07%	8.01%	-0.94%	-1.78%	-1.48%	-0.31%	0.02%	-0.01%	0.03%	-0.03%	0.67%	-0.70%
2015	-0.95%	-0.41%	-0.54%	-0.63%	-0.10%	-0.52%	-0.71%	-0.16%	-0.55%	-1.89%	-1.24%	-0.65%

Source: AACC calculations based on Morningstar data.

4. POTENTIAL DATA ISSUES

There are some data-related observations that may affect the results shown above.

Incomplete Data in Morningstar

Some funds in the Morningstar data appear to be missing asset size or returns/alphas and were hence excluded from the analysis.

Alphas appear to be available only for full calendar years from January to December. Funds that started during 2014 or 2015 or that dissolved before a full calendar year elapsed are therefore missing data. But we have identified a number of funds that lacked data despite existing for at least one calendar year.

In order to gain a complete picture of relative performance of broker-dealer sold funds, it is important to access and analyze data for mutual funds that are no longer active. If the missing data are concentrated among obsolete funds and/or among one classification of funds (broker-sold or direct-sold), this can affect the results shown above and the conclusions derived from them.

The impact of these missing data cannot be measured until the net asset size of the missing funds and their performance can be established.

Using Share Class Type to Determine Sales Channel

In order to measure relative performance of broker-dealer sold funds to direct-sold funds, this analysis uses the share class type field to classify and analyze the data. The analysis assumes that share classes A, B, C, D, T and Adv use the broker-dealer sales channel.

Many load funds may waive load fees in order to gain access to investors served in a fiduciary setting such as Defined Benefit plan or fee-based advisory account. Some of the assets, returns and alpha for a fund with share class type A currently classified as a broker-dealer sold should therefore be instead classified as direct-sold or neither direct nor broker-sold.⁶

⁶ A recent filing by UBS Financial Services Inc., an investment advisory firm subject to a fiduciary standard, shows that some of its clients were invested in load-funds such a class type A funds. (Form ADV Disclosure Brochure, UBS Financial Services Inc. March 31, 2015 available at http://www.adviserinfo.sec.gov/Iapd/Content/Common/crd_iapd_Brochure.aspx?BRCHR_VRSN_ID=308745.)

APPENDIX. SHARE CLASS TYPE DESCRIPTIONS

This appendix is based on Morningstar Direct.

Shares of a fund may be offered in different classes, corresponding to different shareholder rights and obligations, such as different fee and load charges. Common share classes are A (front-end load), B (deferred fees), C (no sales charge and a relatively high annual 12b-1 fee, such as 1.00%). Multi-class funds hold the same investment portfolio for all classes, and differ only in their surrounding fee structure.

Share Classes Associated with Broker-Sold Funds

Share Class – A

Funds that have lower investment minimums and carry a front-load to pay the advisors' sales commission. Front-load discounts are usually available if the investor meets a higher minimum initial purchase. Also known as 1, I or One. Typically, the maximum front load is between 4% and 5.75%, the maximum deferred load is zero, the maximum 12b-1 fee is between 0 and 50 bps and the investment minimum is \$2,500 or less.

Share Class – Adv

Funds typically purchased through advisors, but generally requiring a higher minimum investment. Also known as Adv or Advisor. Typically, the maximum front load is 0%, the maximum deferred load is 0%, the maximum 12b-1 fee is between 0 and 50 bps, and the investment minimum is \$2,500 or less.

Share Class – B

Funds that have lower investment minimums and carry a deferred-load sales charge, also called a surrender charge. The sales charge is imposed if shares are redeemed before specified time periods, typically within five years. The sales charge decreases with the time invested such that the surrender charge is higher in year one than it is in year five. Also known as 2, II, or Two. Typically, the maximum front load is 0%, the maximum deferred load is between 4% and 5%, the maximum 12b-1 fee is between 75 and 100 bps, and the investment minimum is \$2,500 or less.

Share Class – C

Funds that have lower investment minimums and carry a level-load structure. This sales charge is typically a recurring fee of 1% that is used on an annual basis to compensate advisors. Investment minimums for C- shares tend to be lower than for D-shares. Also known as 3, III, or Three. Typically, the maximum front load is 0% and occasionally 1%, the maximum deferred load is 1% and occasionally 0%, the maximum 12b-1 fees is between 75 and 100 bps, and the investment minimum is \$2,500 or less.

Share Class – D

Funds that have lower investment minimums and carry a level-load structure. This sales charge is typically a recurring fee of 1% that is used on an annual

basis to compensate advisors. Investment minimums for C- shares tend to be lower than for D-shares. Typically, the maximum front load is 0%, the maximum deferred load is 0% and occasionally 1%, the maximum 12b-1 fee is 0% and occasionally between 1 and 50 bps, and the investment minimum is \$2,000 or more.

Share Class – T

Typically, T shares carry lower front-end loads than A shares and are available to investors with larger initial investments. Typically, the maximum front load is 0% and sometimes between 3% and 4.75%, the maximum deferred load is 0%, the maximum 12b-1 fee is sometimes 0% and sometimes between 25 bps and 50 bps, and the investment minimum is \$2,000 or more.

Share Classes Associated with Direct-Sold Funds

Share Class – Inv

Investor share classes can be purchased by individual investors, so there is usually no front or deferred load charged. However, investment minimums may be slightly higher. Also known as Investor or Investment. Typically, the maximum front load is 0%, the maximum deferred load is 0%, the maximum 12b-1 fee is sometimes 0% and sometimes between 1bp and 25 bps, and the investment minimum is \$10,000 or less.

Share Class – No Load

Funds without front- or back-end sales charges. Purchased directly by investors or through advisors. Typically, the maximum front load is 0%, the maximum deferred load is 0%, the maximum 12b-1 fee is between 0 and 100 bps, and the investment minimum is \$2,500 or less.

Share Class – S

S share classes are similar to no-load funds in that there is usually no front or deferred load charged. However, investment minimums may be slightly higher. Typically, the maximum front load is 0%, the maximum deferred load is 0%, the maximum 12b-1 fee is 0%, and the investment minimum is \$2,000 or more.

Share Classes Associated with Neither Broker- Nor Direct-Sold Funds

Share Class – Inst.

Funds typically purchased by large institutional buyers, such as pension plans. Also known as Y, I, Z, X, Inst, Instl. Typically, the maximum front load is 0%, the maximum deferred load is 0%, the maximum 12b-1 fee is 0%, and the investment minimum is \$25,000 or more.

Share Class – Load Waived

Load Waived share classes don't require its investors to pay loads (but 12b-1 fee may still be required). In most cases, mutual fund companies will limit the number of load-waived funds available to only certain investors.

Share Class – M

Typically, M shares carry lower front-end loads than A shares and are available to investors with larger initial investments. Typically, the maximum front load is sometimes 0% and sometimes between 1% and 3.5%, the maximum deferred load is 0%, the maximum 12b-1 fee is sometimes 0% and sometimes between 25 bps and 100 bps, and the investment minimum is \$50,000 or more.

Share Class – N

Typically, N shares are available to investors with larger initial investments. Many also charge a 12b-1 fee. Typically, the maximum front load is 0%, the maximum deferred load is 0%, the maximum 12b-1 fee is between 25 and 50 bps, and the investment minimum is \$50,000 or more.

Share Class – Other

Funds not elsewhere classified. This category contains fewer than 5% of all funds. Also known as most other share class letters. The maximum front load varies, the maximum deferred load varies, the maximum 12b-1 fee varies, and the investment minimum varies.

Share Class – Retirement

Funds available through retirement plans. Purchased by retirement plan participants, usually without any sales loads. Also known as Ret, R, K, and J. Typically, the maximum front load is 0%, the maximum deferred load is 0%, the maximum 12-b1 fee is between 25 and 50 bps, and the investment minimum varies.

DISCLAIMER

The views, opinions, and/or findings contained in this report should not be construed as an official Government position, policy or decision, unless so designated by other documentation issued by the appropriate governmental authority.

We call your attention to the possibility that other professionals may perform procedures concerning the same information or data and reach different findings than Advanced Analytical Consulting Group, Inc. (AACG) for a variety of reasons, including the possibilities that additional or different information or data might be provided to them that was not provided to AACG, that they might perform different procedures than did AACG, or that professional judgments concerning complex, unusual, or poorly documented matters may differ.

This document contains general information only. AACG is not, by means of this document, rendering business, financial, investment, or other professional advice or services. This document is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action. Before making any decision or taking any action, a qualified professional adviser should be consulted. AACG, its affiliates, or related entities shall not be responsible for any loss sustained by any person who relies on this publication.

80

 June 5, 2008  *Property and Casualty News* 

Be Careful What You Wish For: As The Industry Changes, Advisors Need To Be Aware Of The Tradeoffs Their Broker-Dealers May Be Making To Stay Profitable

Copyright 2008 SourceMedia, Inc. All Rights Reserved Financial Planning

June 1, 2008

2250 words

BE CAREFUL WHAT YOU WISH FOR: As the industry changes, advisors need to be aware of the tradeoffs their broker-dealers may be making to stay profitable.

Philip Palaveev

The independent broker-dealer industry is at an important crossroad. How advisors work and how b-ds survive and serve their clients have been in some cases on a collision course. Understanding this conflict is important to independent advisors who want to choose a broker-dealer that truly suits their needs and where they will feel comfortable and confident. While it's true that only some 10% of advisors are looking to change their firms in any given year, this discussion is not only relevant to them. It can be argued that every year advisors "renegotiate" their contract with their firm when they chose to stay.

Often advisors don't know how to fully evaluate a broker-dealer and so they focus in on payout as the overriding factor. But as payouts continue to rise and profit margins for b-ds continue to shrink, advisors have to realize that some b-ds may be making their profits in ways that conflict with the interests of advisors and their clients.

Advisors probably know much of this already but choose to consider the economic challenges of the broker-dealers as "not my problem." Maybe it is their problem, though, because the economics often define the relationship. Knowing how a broker-dealer makes money can help advisors understand the trade-off between the payout they receive and the potential conflicts of interest they are willing to live with.

DISAPPEARING "REP"

The conflict all has to do with the shifting role of the advisor. Traditionally broker-dealers were created to supply financial products to registered representatives, who then sold them to consumers. The profits from the products were the center of the economic relationship and the role of the broker-dealer was to maximize the throughput. The distribution piece of the equation didn't have to do much more than break even.

Today, however, the center of the relationship has shifted from the product to the skills of the advisor. It's rare for advisors to describe themselves as registered representatives of a broker-dealer anymore. Indeed the term "rep" is quickly disappearing from the vocabulary of even broker-dealer executives themselves. This shift in terminology is indicative of the fundamental shift in the role of affiliated advisors. Today's advisors are in the catbird seat: they control the relationship with the client and their advice is not so easily replaced. As a result, they have gained substantial bargaining power with broker-dealers. In addition, advisors have become a limited resource: There aren't enough of them to meet the rising consumer demand, so b-ds are having to compete ever more vigorously to recruit them. Consequently, payout ratios are steadily rising with no end in sight.

This puts broker-dealers in dicey position. Even in good times, their profit margins were never spectacular, and the profitability picture has been far from stellar for the past several years. The independent broker-dealer industry saw losses of 1% to 0.5% from 2001 to 2003, followed by a slight improvement in 2004 and 2005 when the average broker-dealer profits reached 1.5% industry-wide. But that profit has eroded recently, according to Moss Adams's 2007 FSI Financial Performance Survey. It's been hard to manage the bottom line with very competitive payouts. At the same time service expectations have increased and with them the cost of being a broker-dealer. Based on Moss Adams data, the number of compliance employees per dollar of business has doubled. The number of staff in marketing and practice management has also doubled, and technology expenses as percent of revenue have increased too. Considering these increases in cost and an environment where payouts over 90% are frequently advertised, it may appear that a broker-dealer has to provide all of these services for less than 10% of the revenue.

In reality, broker-dealer expenses are typically higher than 10%, but also the effective payout is not 90%. The fact is that not all advisors qualify for the 90%-plus payout. In fact, as a percent of commissionable revenue, broker-dealers pay out about 82% of the total revenue, for an 18% gross margin, which is largely eaten up by roughly equivalent operating expenses (approximately 12% in operating expenses and 6% in other direct expenses such as clearing and other costs). It would clearly be difficult for most broker-dealers therefore to pay out 90%+ to their advisors. But b-ds don't generally have to make such rich payouts primarily because the average advisor generates \$126,000 in gross dealer concession (broker-dealer revenue). Very few broker-dealers have average revenue per advisor of \$250,000 or more. Indeed it's ironic that b-ds tend to chase only big producers, when the best advisors for them may be mid-size. After all, the backbone of the gross margin profitability for a broker-dealer comes from the advisors who produce \$250,000 to \$750,000. These advisors haven't yet entered into the 90%+ payout range.

WHERE THE MONEY COMES FROM

B-ds also have some "hidden profit centers" and this is where it's important for advisors to pay attention to b-d economics. Approximately 6% to 8% of the total average revenue of independent broker-dealers represents a precious cache of high-margin services and products. These non-commissionable revenues include any marketing fees from mutual fund and investment management funds plus internal asset management programs. Fees from these sources can potentially represent a conflict of interest because b-ds have incentive to promote such funds and programs even if they aren't in the long-term interest of clients. Rep fees and charges, interest on cash in the client accounts, and markups on anything from ticket charges to E&O insurance and software are other unrevealed sources of income, which force advisors and clients to unwittingly pay extra for services.

Of course, not all broker-dealers do this. Some broker-dealers have accepted the shift from a "distribution" model to an "advisory" model and have retooled themselves. Very large, well-managed broker-dealers can make up the margin difference by finding a balance between using scale and leverage with suppliers to create a profit of 5% to 7%. A very large firm can negotiate the lowest cost from its vendors, including clearing partners, mutual funds and managers. A smaller one will not have this leverage. So if an advisor is looking for a smaller firm that can provide the same payout as LPL, the advisor has to realize that the firm will either have to be willing to live with lower or no profits or will have developed some other sources of profitability that may smack of conflict of interest.

WHAT ADVISORS NEED TO KNOW ABOUT B-DS

If advisors were to consider the economics of broker-dealers in their choices, they would find that the independent broker-dealers can be classified in the following categories:

* The large and efficient firm: This firm provides infrastructure, technology and streamlined processes. It has a high payout because it's very large and has scale. It recruits large practices and rewards them with even higher payouts. It has a lot of resources, but most will be of the "self-help" kind-manuals instead of consultants. The firm may use its scale to gain additional revenue from the vendors but generally the platforms are open with many choices and very competitive pricing. LPL, most of all, dominates this category. Other firms like Securities America, Royal Alliance and NFP fall into this category as well.

* The small firm: This is a firm that provides cultural affiliation first and only secondly tries to match the capabilities of bigger competitors. Executives are more accessible: Advisors can call the broker-dealer's president and chat with him on the phone. The firm ramps up its capabilities through the "borrowed" scale of the clearing firm. But the firm usually lacks the ability to customize software programs and platforms for its advisors and must rely on whatever the clearing firm offers. This b-d is not in a position to offer high payouts, but it will probably yield to pressure from the recruiting race and try. In the long term, the higher the payout, the greater the chance that the firm will be sold to a larger distribution-oriented competitor. Firms such as Geneos and Triad advisors have created the culture of a small firm with high-end service.

* The branded firm: This firm has a consumer brand that attracts clients and helps the advisor grow their practice faster. The brand requires resources to market and establish its brand and thus payouts are much lower than at other firms. Raymond James is a great (and perhaps the only) example of this category among independent broker-dealers. The future could bring more entries into this category as it's an open competitive space, but that may be prohibitive for independent broker-dealers given the resources required to create a consumer brand.

* The unique business model firm: This firm offers an affiliation or pricing model no one else is providing. A great example is Cambridge Investment Research, which supports hybrid firms that are both RIAs and affiliated. Although other firms have subsequently adopted this service (LPL just announced a new platform for hybrid RIAs), Cambridge has successfully used this unique positioning to become a large and efficient firm too.

* The specialized firm: Several broker-dealers target specific markets such as CPAs, credit unions, banks and others. The advantage of such firms is their specialized knowledge, training and compliance resources. The payouts are lower, and the technology is standard and borrowed from the clearing firm. The profits come the old-fashioned way: total revenue minus payouts and expenses. The niche protects the firm from the competition. First Global, Genworth and HD Vest for example dominate the CPA space. UWest (part of LPL) and PrimeVest (ING) are the largest bank broker-dealers.

* The distribution firm: Many financial conglomerates still have broker-dealers that recruit and retain advisors to distribute the company product, although usually not exclusively. Payout is very high on "preferred" products. Operations at such firms are probably a little slower and product capabilities a little more restrictive. The compliance department is larger since the distribution mentality creates some risk in itself. The firm has many production incentives and "president's circles" and incentive trips. The cultural affiliation may be strong, but only for advisors who grew up in the system. The notion of product distribution may extend beyond the insurance product or proprietary funds. For example, proprietary investment advisory platforms have created a new type of "product" for many firms by supplying an internal and highly profitable source of revenue.

* The bad choice firm: This is everybody else. Not large enough, not efficient enough, not small enough, not specialized enough, but trying to compete on all of these parameters. It's a firm without much of an identity that tries to be all things to all advisors. It tries to compete on payout and struggles to find a way to create either scale or differentiation. Unfortunately, this probably describes too many broker-dealers.

Note that of all the categories, only the large firms and the distribution firms are in a position to compete on payout. All other categories need to rely on some unique differentiating factor. If advisors insist on the highest payout and they are not going to a very large firm, they are in essence getting a distribution firm.

The negotiating power advisors have today is unprecedented, but they have to use it carefully and direct it at what they really want or need from a broker-dealer. Advisors who focus only on the payout may not realize that making a decision that way could ultimately come back to them in the form of potential conflicts of interest and higher undeclared fees. Most advisors are shifting away from the distribution model, but elements of that model are still prevalent in the b-d industry. To use a cooking metaphor: Advisors are no longer just using the company recipe-book and can create any meal they want using the best ingredients available. But if they want to receive premium ingredients, they have to either accept premium prices in the form of a somewhat lower payout or negotiate the price (payout) and live with a lot of house brands, which are not necessarily premium ingredients.

The rising negotiating power of advisors gives them more choices but also puts more responsibility on advisors. Each advisor has to think in terms of what do they expect and what do they need from their broker-dealer in order to develop their practice. If the answer is simply "the highest payout" then advisors should be prepared to accept a more streamlined service model. If the answer is value added solutions then some of the cost of such solutions will affect the payout. Demanding a high payout and maximum service simply recreates the negative aspects of the distribution approach the industry is trying to leave behind.

Philip Palaveev is the president of Fusion Advisor Network, a \$50 million franchise of independent advisors that provides business management and collective bargaining services to its members. Prior to joining Fusion, Palaveev was a principal of accounting and consulting firm Moss Adams LLP in Portland, Ore. (c) 2008 Financial Planning and SourceMedia, Inc. All Rights Reserved. <http://www.Financial-Planning.com> (<http://www.Financial-Planning.com>) <http://www.sourcemediacom.com> (<http://www.sourcemediacom.com>)

June 2, 2008

Copyright © 2008 LexisNexis, a division of Reed Elsevier Inc. All rights reserved. (<http://www.lexis-nexis.com/lnc/about/copyrt.html>)

Terms and Conditions (<http://www.lexis-nexis.com/terms/general>) Privacy Policy (<http://www.lexis-nexis.com/terms/privacy>)

© 2018 InsuranceNewsNet.com, Inc. All rights reserved.

81

Mutual fund investors: sharp enough?

Alan R. Palmiter
Professor of Law, Wake Forest University

Ahmed E. Taha
Professor of Law, Wake Forest University

Abstract

Who are mutual fund investors? The answer is critical to regulatory policy. The mutual fund industry portrays fund investors as diligent, fairly sophisticated, and guided by professional financial advisors. The SEC paints a more cautious portrait of fund investors, though touts improved disclosure by the fund industry as a sufficient antidote. However, an extensive academic literature finds that fund investors are unaware of the basics of their funds, pay insufficient attention to fund costs, and chase past performance despite little evidence that high past fund returns predict future returns. These findings suggest that policymakers should rethink current regulatory policy. Disclosure may not be enough.

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/46557869>

Mutual Fund Investors: Sharp Enough?

Article · January 2008

Source: RePEc

CITATIONS

0

READS

98

2 authors, including:



Alan R. Palmiter
Wake Forest University

23 PUBLICATIONS **184 CITATIONS**

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:

Project

Corporate Triplespeak: Response of IOUs to Clean Power Plan View project

Mutual fund investors: sharp enough?

The U.S. mutual fund industry is subject to the decisions and discipline of dispersed and diverse mutual fund investors. Their investment knowledge and sophistication are critical to the effective functioning of the fund market. Unlike the markets for publicly-traded securities, the fund market lacks mechanisms – like short-selling – to ensure informational efficiency. This article examines the profiles of mutual fund investors presented alternatively by the mutual fund industry, the Securities and Exchange Commission (SEC), and by an extensive academic literature.

Mutual fund ownership

Mutual funds have transformed from being the 'small investor's diversified portfolio' to 'every person's retirement account.' Over the past two decades, U.S. retirement assets have shifted from defined-benefit (pension) plans to defined-contribution plans and IRAs. As of 2007, U.S. retirement assets totalled U.S.\$17.8 trillion, with U.S.\$4.5 trillion in defined-contribution plans and U.S.\$4.8 trillion in IRAs. About half of the defined-contribution assets (U.S.\$2.4 trillion) and IRA assets (\$2.3 trillion) are invested in mutual funds. In all, the total share of retirement assets held by mutual funds increased from 8% in 1992 to 26% in 2007 [ICI Research Fundamentals (2008)].

Mutual fund ownership is widespread. Today, almost 51 million American households (or 44%) own mutual funds, far more than those holding individual stocks and bonds [ICI Ownership Trends (2007)]. The median fund-owning household has a modest income of U.S.\$74,000 and financial assets of U.S.\$175,000, of which U.S.\$100,000 is invested in mutual funds [ICI Fund Shareholders (2008), ICI Fact Book (2008)]. Mutual fund holdings total almost U.S.\$12 trillion. Half of these holdings are in equity funds (50%), with most of the rest divided between money market funds (29%) and bond funds (15%) [ICI Investing Trends (2008)].

Industry's portrait of fund investors: sophisticated and informed

The Investment Company Institute (ICI) is the national trade association for the mutual fund industry. It has consistently maintained that the industry is highly competitive, with fund choices, services, and fees set in a robust market of fairly sophisticated and informed investors.

ICI survey of (some) fund investors – in 2006, the ICI surveyed 737 investors who had bought a new stock, bond, or hybrid fund outside of an employer-sponsored retirement plan [ICI Investor Preferences (2006)]. The survey found that the information investors most commonly consider before investing are a fund's fees and expenses (74%), its historical performance (69%), and its risks (61%). In addition, the survey found that fund investors consult an average of three information sources before investing, the most common being a professional financial advisor (73%).

The ICI survey paints a very comforting portrait of mutual fund investors: investors consider important information before investing and are usually assisted by a professional. But the survey may be of limited value. In addition to relying on investors to accurately self-report their behavior in face-to-face interviews, the survey was limited to investors who own funds outside of employer-sponsored plans. Such investors are likely to be more experienced and financially sophisticated than the typical investor. In addition, most of those surveyed report using professional financial advisors, which is a resource often unavailable in employer-sponsored plans.

ICI statements about fund investors – the ICI's profile of mutual fund investors is also reflected in other ICI publications, comment letters to the SEC regarding proposed rules, and statements by its officials. The ICI frequently asserts that investors search for funds with low fees and expenses. ICI publications regularly state that this imposes a strong market discipline on mutual funds. ICI officials echo this view in speeches and interviews. As evidence, the ICI points out that investors are heavily invested in lower-cost funds. An ICI study of mutual fund industry competition concluded that "more than three-quarters of stock and bond fund assets are invested in funds charging below-average operational and management expenses" [ICI Competition (2006)]. Also, another ICI study found that the average fund fees and expenses that investors paid in 2005 were at their lowest level in more than 25 years [ICI Fees and Expenses (2006)].

Nonetheless, while trumpeting investor demand for lower-cost funds, the ICI has been careful not to encourage investors to invest in such funds. For example, an online mutual fund primer sponsored by the ICI Education Foundation and the National Urban League points out that annual stock market returns have averaged 11% since 1926, yet is silent about the effect of expenses, fees, and loads on fund returns.

ICI views on investor access to information – ICI officials have also credited as a source of market discipline the great amount of fund information available to investors. They claim that this information allows investors to compare funds, forcing funds to compete for investors. Interestingly, however, the ICI also believes that most investors are receiving too much information about funds. Thus, the ICI has supported steps taken by the SEC to reduce, summarize, and standardize the information provided in fund prospectuses and to create even more basic disclosure documents, such as fund profiles. In addition, the ICI has supported SEC initiatives to simplify the language used in disclosure documents.

SEC's portrait of fund investors: capable (with some help)

The SEC, though less sanguine than the fund industry, also paints a portrait of fund investors that implies a well-functioning mutual fund market. The SEC shares the industry view that fund investors con-

Mutual fund investors: sharp enough?

sider a variety of important information before investing. The SEC also agrees that fund investors can benefit from fund information being presented in a concise, readable, and standardized form. Unlike the industry, however, the SEC has specific concerns about investor sophistication, worrying that fund investors pay too little attention to fund fees and expenses and too much attention to past returns. The agency's solution, however, is generally improved disclosure.

As then SEC Chair Arthur Levitt told Congress, "The Commission should not be the arbiter of the appropriate level of fund fees. Whether fund fees are too high or too low is a question that we believe must be answered by competition in the marketplace, not by government intervention."

SEC regulation of fund disclosure – by law, prospective fund investors must receive a prospectus before or when they invest. The SEC views the fund prospectus as the primary source of information for investors choosing a fund. In addition, the SEC requires that fund advertisements explicitly direct investors to the prospectus by warning "the prospectus should be read carefully before investing." The prospectus, aimed at the 'typical investor,' must include information about the fund's investment objectives or goals, principal investment strategies, principal risks, 10-year returns, returns of an appropriate market index, fees and expenses, fund management, fund pricing, purchase and redemption, fund marketing (and loads and 12b-1 fees), and financial highlights.

Except for disclosing the returns of an 'appropriate' market index, the prospectus need not provide any information on performance, loads, fees, expenses, or portfolio turnover of comparable funds. The SEC assumes fund investors, left to their own devices, will determine how a fund compares to other funds. Likewise, the prospectus need not educate investors about diversification, long-term returns of various asset classes (such as stocks or bonds), or the relationship of risks and returns. By not requiring such information in the prospectus, the SEC implicitly assumes that most investors already know or do not need this information.

In its most recent effort to improve disclosure to fund investors, the SEC proposed in November 2007 to "provide investors with streamlined disclosure of key mutual fund information at the front of the statutory prospectus, in a standardized order that facilitates comparisons across funds." The proposed summary, written in 'plain English' and 3-4 pages long, would contain information in a fixed order, including information not currently required in the current prospectus such as risk/return bar chart, volatility of returns, top ten portfolio holdings, identification of portfolio managers, and financial intermediary compensation.

The SEC also assumes the format in which information is presented is important to investors. It believes fund investors are more likely

to understand clear, concise, and standardized disclosure in plain English or in graphical or tabular form. Thus, for example, the prospectus must contain a standardized graph and table of the fund's historical returns.

Nonetheless, despite its awareness of the difficulties investors have in processing information, the SEC has not ensured that it is really helping them understand their fund investment choices. In its 1998 prospectus disclosure rulemaking intended to improve the communication of important information about funds to investors, the SEC acknowledged an ICI survey finding that about half of fund investors do not read the prospectus before investing, but the SEC has not conducted a follow-up study on whether fund investors are reading or understanding the new streamlined prospectuses it mandated.

SEC regulation of fund advertising – mutual fund advertisements are also an important source of information for investors. Investors invest more in heavily advertised funds. The SEC's rules regarding fund advertisements, like its rules regarding prospectus disclosures, reflect its view of fund investors. One recurring theme of the advertising rules is that investors must be reminded to consider certain important information when choosing a fund. Fund advertisements must explicitly advise investors "to consider the investment objectives, risks, and charges and expenses of the investment company carefully before investing."

The SEC's advertising rules also acknowledge that investors may overweight past performance, which mutual funds with strong records routinely tout despite little evidence that past returns predict future returns. Thus, it requires advertisements of past returns to include a boilerplate warning that "past performance does not guarantee future results." In addition, any performance information in advertisements must be standardized. For example, the SEC prescribes how returns must be calculated and specifies that returns must be presented for one-, five- and ten-year time periods.

As with its prospectus regulations, the SEC's advertising regulations are based largely on the agency's unsupported perceptions of the impact of its mandates. For example, although the SEC requires that advertisements warn that past returns do not guarantee future returns, there is no indication that the SEC has ever examined whether these warnings are effective.

Academic literature's profile of fund investors: mostly clueless

The academic literature finds fund investors ill-prepared to choose among funds. Contradicting the industry's and regulator's profiles of relatively sophisticated fund investors, an extensive body of research of actual investor behavior – conducted primarily by finance professors – reveals that most fund investors are unsophisticated and uninformed. This finance literature finds that investors

are mostly ignorant of the basic characteristics of the funds they own. They are unaware of their funds' riskiness and pay insufficient attention to fund costs. In addition, they regularly buy funds based on strong past performance, despite the scant correlation between high past and future returns.

In short, mutual fund investors are unable to perform the disciplining role that the industry says they do and that the SEC believes they can do with some assistance. This is unsurprising. By definition, mutual fund investors have chosen to delegate management of their investment to a professional. But fund investors still must choose in which funds to invest. The academic literature suggests that they need much more help to do this than they are receiving now.

Investors are ignorant of basic fund characteristics – Capon et al. (1996) conducted a survey of fund investors and found that they are “in general uninformed regarding the nature of their investments.” Of those surveyed, 72% did not know if their primary fund invests in domestic or international securities, and 75% did not know whether the fund invests in equity or fixed income securities. Similarly, a survey of 2,000 mutual fund investors, sponsored by the SEC and the Office of the Comptroller of the Currency, found that 57% of them did not know the expense ratio of their primary fund even when they purchased it [Alexander et al. (1998)].

Investors are inattentive to risk – fund investors are generally not aware of fund risks and do not take them seriously. A number of academic studies of investor behavior are at odds with the conclusions of the ICI Investor Preferences survey (2006), in which most respondents (61%) reported they review a fund's risks before investing, making ‘fund risks’ the third most reviewed information by fund investors. Academic studies show fund investors consistently do not understand and are indifferent to risk. Wilcox (2003) asked fund investors to choose among hypothetical stock funds differing in up to six characteristics: the fund company's name, the fund's load, the fund's annual management fee, the fund's returns during the previous year, the fund's average annual returns during the previous 10 years, and the fund's beta. He found that a fund's beta was the least important characteristic to investors. Del Guercio and Tkac (2002) found that surveyed investors overwhelmingly acknowledge not using standard measures of risk – such as beta or standard deviation of fund returns – in evaluating a fund. In addition, they found that investors do not pay attention to whether a fund minimizes even its diversifiable risk.

Even though investors personally do not carefully consider a fund's risk, many investors say that they use published fund ratings or rankings – such as those from Morningstar – that often incorporate some measure of risk. The ICI survey, for example, found that 35% of investors review a fund's “rating from a mutual fund rating service” before purchasing the fund, and 19% call it “very important” to their

final decision to invest. But this reliance on third-party advice does not appear to markedly increase investors' sensitivity to fund risk. Although there is a positive relationship between flow and funds' risk-adjusted returns, Del Guercio and Tkac found that it is not as strong as the positive relationship between flow and non-adjusted returns.

Investors pay insufficient attention to fund expenses – studies generally show that operating expenses, paid through management fees, matter little to many investors.

Surveys of fund investors

Surveys of fund investors have found that they pay little attention to fund fees and expenses. For example, in a survey sponsored by the SEC and Office of the Comptroller, only 19% of fund investors could give even an estimate of the expense ratio of their largest mutual fund holding, and only 43% claimed to have known it even at the time they first purchased the fund [Alexander et al. (1998)]. Similarly, Wallison and Litan (2007) surveyed mutual fund investors and asked them to rate (on a scale of one to ten) how informed they were about their own mutual fund investments. Only 21% of respondents rated their knowledge highly (8 or above), and less than 10% of even this self-described knowledgeable subgroup “knew even approximately what they were paying as an advisory fee.” Also, Capon et al. (1996) asked mutual fund investors to rate the importance of nine particular factors in choosing a mutual fund. The ratings could range from one (not at all important) to five (extremely important). The respondents gave management fees an average rating of only 2.28, fifth among the nine factors.

These results stand in stark contrast to those of the ICI survey in which 74% of those surveyed claimed to review fund fees and expenses before investing. One reason for the discrepancy may be differences in the questions asked. The ICI survey asked whether the investor “reviewed” a fund's fees and expenses before investing, while the Capon et al. survey, for example, asked how “important” the fund's fees were to the investor in choosing a fund. Investors could review a fund's fees before investing, yet not give it much weight. This difference, however, would not reconcile the ICI survey with the SEC and Office of the Comptroller survey, in which less than half of investors reported knowing their largest fund's expenses when they purchased the fund.

Studies of investor behavior

Although many investors seek funds with low expenses, studies reach mixed conclusions regarding whether funds with lower expenses receive more flow overall. The Barber et al. (2005) study of diversified U.S. equity mutual funds from 1970-1999 found “at best, no relation between operating expenses and flows and, at worse, a perverse positive relation between expenses and flows for large funds.” In addition, they examined the mutual fund purchases

and sales made by customers of a large discount broker for 1991-1996, identifying those instances when a fund purchase followed a sale within three weeks. They found that, for these instances, the operating expenses of the purchased funds were generally higher than those of the funds that were sold, indicating that investors do not switch funds to reduce the expenses they are paying.

In contrast, the Sirri and Tufano (1998) study of equity mutual funds from 1971-1990 found that reductions in expenses have a strong positive effect on flow, yet they also found that increases in expenses strangely have a marginally significant positive effect on flow as well. In addition, Zhang (2007) and Coates and Hubbard (2007) found that investor demand for mutual funds is elastic with respect to fund expense ratios and Zhang also that the elasticity has been increasing recently. However, Coates and Hubbard (2007) note that recent studies have failed to find consistently that fund expense ratios have declined in recent decades, even though the industry enjoys large economies of scale and the size of funds and fund families has grown dramatically. For example, an SEC study (2000) found that the weighted average expense ratio for equity and bond funds rose from 0.73 in 1979 to 0.94 in 1999. Thus, despite the ease with which fund investors may move from one fund to another, investors seemed not to have had a great disciplining effect on expenses.

Evidence from index funds

The growth of low-cost index funds over the past 30 years provides evidence that a subset of investors is sensitive to fund expenses. Index funds' popularity, however, has plateaued. Index funds first became available in 1976, and by 1999 grew to hold 10% of equity mutual fund assets. Index funds' market share, however, has remained stable since then [Bogle (2007)]. In addition, the success of high-cost index funds indicates that many investors continue to disregard fund expenses. Although a rational investor might buy a high-cost, actively-managed fund hoping the fund manager's stock picking skills will compensate for the higher expenses, a high-expense index fund makes little sense. By definition, an index fund manager's stock picking skill is irrelevant. A rational index fund investor should buy the index fund with the lowest expense ratio, because all index funds are holding essentially the same securities.

Despite this, Elton et al. (2004) found that a large amount of new cash flows into S&P 500 index funds from 1996-2001 went "to the poorest-performing [high-cost index] funds." Also, new index funds that entered the market during the period had higher expense ratios (0.77 average, 2.00 maximum) than the average index fund (0.44). And the highest-cost decile of index funds grew almost twice as much as the lowest-cost decile. Similarly, Hortacsu and Syverson (2004) found that average expense ratios for S&P 500 index funds rose from 0.27% in 1995 to 0.32% in 2000, with high-expense S&P 500 funds (highest quintile) nearly tripling their market share.

The fund industry has attempted to explain the success of high-expense index funds. A recent ICI publication argues that differences across index funds, such as the size of fund assets and average account balances, may explain much of the differences in their expense ratios [Collins (2005)]. Although smaller funds may have higher costs, it does not explain why investors would choose such funds when identical, lower-cost alternatives exist. In the same publication, the ICI also argues that fund investors may choose high-cost index funds because such funds offer better services to investors. However, Elton et al. (2004) found no relationship between fund families' service quality and investors' choice of index funds. They also found no evidence that investors' purchases were concentrated in the largest fund families, which should be better able to provide services than smaller families.

Investors (increasingly) pay attention to loads

Many fund investors, although generally insensitive to fund operating expenses, increasingly pay attention to fund loads, which are becoming less common.

Studies on loads – do loads increase or decrease fund flows?

On the one hand, investors should be averse to paying additional fees, so one would expect a negative relationship between loads and flow. In addition, because loads are one-time fees that are deducted directly from an investor's individual account they are much more salient than operating expenses. Indeed, in an experiment requiring investors to choose among hypothetical stock funds, Wilcox (2003) found that investors were more averse to front-end loads than to high expense ratios. On the other hand, load proceeds are used to market funds, so additional marketing might offset investor aversion to loads. Studies have come to differing conclusions regarding the net effect of loads on fund flow. Sirri and Tufano (1998) found that for equity funds from 1971-1990 there was no relationship between loads and flow; investors' aversion to loads seemed to roughly offset the effect of additional marketing. More recently, however, Barber et al. (2005) found a negative relationship between loads and flow for equity funds from 1970-99.

Although many fund investors are sensitive to loads, many other investors clearly pay little attention to them. Capon et al. (1996) found in a survey of households that invest in mutual funds that 39% of respondents stated they did not even know if their primary fund has a load. In addition, Wilcox's experiment (2003) found the fund's load was only the fourth most important factor (of six) to the investors – ahead of the fund's management fee and beta, but behind the fund's return over the past ten years, the fund's return over the past year, and the fund company's name.

Rule 12b-1 fees – investors pay less attention to 12b-1 fees, the ongoing marketing fees deducted over time from the fund's assets.

Mutual fund investors: sharp enough?

Like loads, 12b-1 fees are used to pay broker commissions and other marketing costs, but unlike loads, they are not one-time fees deducted directly from investors' individual accounts. Overall, studies indicate that investors are less averse to 12b-1 fees than to other types of marketing fees. Barber et al. (2005) examined the effect of 12b-1 fees flows into diversified U.S. equity mutual funds from 1993-1999. They found that 12b-1 fees rose from 0.14% in 1993 to 0.20% in 1999. Also, unlike their findings for loads, they found a positive relationship between 12b-1 fees and flows, indicating that the marketing benefits from 12b-1 fees outweigh investors' aversion to paying them. In addition, Elton et al. (2004) found that 12b-1 fees did not have a significant net effect on flow into S&P 500 index funds from 1996-2001. In all, investors may have become sensitive to loads, but remain insensitive to smaller, ongoing 12b-1 fees. That is, fund investors notice more when struck by a club, than when slowly bled to death.

Investors chase past returns

Although fund investors pay insufficient attention to fund risk and operating expenses, they pay great attention to historical returns. Studies find this to be the most prominent trait of fund investors. Investor surveys and experiments have uniformly found the importance of a fund's past returns to investors choosing among funds. For example, Wilcox (2003) found that investors choosing between hypothetical mutual funds treated returns over the past ten years and over the past year as the two most important factors. Also, Capon et al. (1996) found that a fund's "investment performance track record" was the most important factor.

Numerous other studies have confirmed that investors flock to mutual funds with the highest past returns. For example, Del Guercio and Tkac (2002) examined flow into a large sample of equity mutual funds and found that a fund's past return has a strong positive effect on fund flow. This positive relationship was strongest for funds with the highest returns. Similarly, Sirri and Tufano's study (1998) of flow into equity funds had similar findings.

Although investors chase the highest returns, there is little reason for them to do so. Despite numerous studies, "within the finance literature there is [only] weak and controversial evidence that past performance has much, if any, predictive ability for future returns" [Wilcox (2003)]. "While some controversial evidence of persistence [of mutual fund returns] does exist ... it is concentrated in low-liquidity sectors or at shorter horizons" [Berk and Green (2004)].

Financial advisors provide little help

The ICI emphasizes that many fund investors consult professional financial advisors. Little evidence exists, however, that these advisors provide investors tangible benefits, and investors pay a high price for using them.

Use of financial advisors – According to the ICI survey (2006), 73% of investors surveyed consulted a professional financial advisor before buying a fund. Similarly, Capon et al.'s survey (1996) of households that invest in mutual funds found that – of nine sources of information – commission-based financial advisors were the third most important source. At first glance, the widespread use of financial advisors by mutual fund investors is encouraging. Although most investors are not financially sophisticated, many are receiving professional advice that might help them overcome their limitations. Before reaching such a conclusion, however, advisors' behavior must be understood.

Survey of financial advisors – Jones et al. (2005) surveyed 530 professional financial advisors and asked them the importance that 14 fund characteristics played in their recommendations to clients. First, the good news. Financial advisors ranked the fund's objective as the second most important factor and fund risk as the third most important factor in their recommendations, giving these two factors much greater emphasis than investors.

Now the bad news. Advisors ranked a fund's expenses as only the eighth most important factor among the fourteen examined. Also, unsurprisingly, marketing costs (loads and 12b-1 fees) were the least important factor to advisors – even less important than to fund investors themselves. In addition, like investors, financial advisors place great emphasis on the fund's past returns. Advisors in the survey ranked comparable performance as the most important factor and absolute performance as the sixth. Thus, financial advisors give high priority to funds that have performed well in the past, despite little evidence that strong past performance predicts strong future performance.

Advisors' effects on actual investor behavior – unfortunately, brokers direct investors to funds that offer brokers more compensation. Bergstresser et al. (2007) compared fund choices from 1996-2004 by fund investors who bought through direct channels and by those who bought through brokers. They found that brokers directed investors toward funds with higher front-end loads and 12b-1 fees. Also, the larger a fund's load and 12b-1 fee, the more flow it received from broker-directed investors. As a result, investors pay a steep price for using brokers. Bergstresser et al. estimated that, in 2002, investors paid up to U.S.\$3.6 billion in front-end loads, U.S.\$2.8 billion in back-end loads, and U.S.\$8.8 billion in 12b-1 fees. However, they found little evidence that fund investors receive commensurate benefits from brokers. For example, brokers do not reduce investors' chasing of past performance; investors who buy funds through brokers are as likely to chase 'hot' funds as are investors who buy through direct channels. In fact, funds purchased through brokers tend to underperform funds that investors buy directly, even before taking into account the higher distribution-related expenses (such as loads) of broker channel funds. This

Mutual fund investors: sharp enough?

underperformance (U.S.\$4.6 billion in 2004 alone) exists in both absolute returns and in risk-adjusted returns.

They also found no evidence that “brokers provide superior asset allocation advice that helps their investors time the market.” The asset allocation choices made through broker-channel funds did not earn higher absolute or risk-adjusted returns than funds bought through direct channels. In theory, investors who buy funds through brokers, thus paying more for distribution, might still benefit if broker channel funds have lower non-distribution expenses (such as management fees). That is, paying for fund marketing to attract new investors and bring in more assets might be worthwhile if, as a result, the fund gained economies of scale and passed them through to existing investors. Bergstresser et al., however, did not find that funds bought through brokers have lower non-distribution expenses – so investors do not benefit in this way either.

Advertising does not benefit investors

In addition to financial advisors, investors pay attention to mutual fund advertisements. Unfortunately, these advertisements do not help investors make better fund choices. Instead, they exploit the tendency of investors to chase past returns. Mutual funds advertise because it works. Capon et al.'s survey (1996) found that advertising was the second most important source of information for investors. In addition, Jain and Wu (2000) found that equity funds that advertised in Barron's or Money magazine experienced approximately 20% greater flows than similar funds that did not advertise. Also, more heavily advertised funds received even greater flows.

Advertising benefits mutual fund companies by increasing flows and thus management fees, which are generally based on assets in the fund. Because investors chase past returns, fund companies often advertise high past performance. For example, Mullainathan and Shleifer (2005) found that past returns were mentioned, on average, in 62% of equity fund ads in Money and in 59% of equity fund ads in BusinessWeek. In addition, Jain and Wu (2000) found that advertised funds outperformed non-advertised funds with the same investment objective by an average of 6% over the twelve months prior to being advertised.

Although advertising benefits fund companies, little evidence exists that it also benefits investors. Jain and Wu found that, after being advertised, funds actually tend to underperform the same benchmarks that they beat before being advertised.

Fixing a dysfunctional market

Fund investors erroneously focus on past returns, while paying insufficient attention to fund risks and ongoing costs. The disclosure and warnings required by the SEC, as well as the advice of financial advisors, are not sufficient to help investors overcome

their limitations. What should the SEC do to improve the functioning of the mutual fund market?

Recognize failure of current disclosure system – the SEC requires fund advertisements to advise investors to read the prospectus and requires that critical information in the prospectus be presented in plain English and in standardized formats. However, few investors refer to the prospectus. As a result, the SEC disclosure regime creates only the illusion of informing investors.

Highlight fund costs – the SEC is considering a prospectus summary that, among other things, highlights expenses in a simplified and standardized tabular form. This is salutary – investors are more likely to use clear, simplified information. For example, Choi et al. (2006) found that investors given a one-page expense summary sheet with the prospectus are more likely to pay attention to expenses and choose lower-expense funds than investors given only the prospectus. However, the current proposal does not call for comparability data. Cox and Payne (2005) have urged that mutual funds disclose how their expense ratios compare to those of comparable funds. They argue that this data is necessary to help investors overcome an evaluability problem. That is, knowing a fund's expense ratio is helpful only if the investor can also readily compare it to that of other funds.

The mutual fund industry has criticized efforts to require comparative data on expense ratios, arguing that it is difficult to determine which funds are comparable to others. However, this argument is a red herring. Investors choosing among mutual funds must already compare funds, and the SEC or mutual funds themselves are unquestionably better able than most investors to determine which funds are comparable. In addition, mutual funds must already make comparability determinations: the prospectus must compare a fund's historical returns to that of a relevant market index.

Nonetheless, leading a horse to water is no assurance it will drink. Even if cost information were displayed more prominently and presented on a comparative basis, fund investors would likely still not give it sufficient weight given their obsession (and that of an obliging industry) with past returns. Investors who believe that a strong-performing fund will continue to outperform its peers are likely to minimize the importance of expenses. Thus, providing information to fund investors on actual persistence of returns may be a key to them paying more attention to fund costs.

Discourage reliance on past performance – the SEC's mandated warnings greatly understate the dangers of chasing past returns. For example, performance advertisements must state words to the effect that “past performance does not guarantee future results [and] current performance may be lower or higher than the performance data quoted.” Such a warning is actually misleading. It sug-

gests that there is a positive relationship between past and future high returns, but just not a 'guaranteed relationship.' Investors, however, need to understand that past returns are not a very useful predictor at all and that strong past performance is generally just a matter of chance [Barras et al. (2008)]. The SEC's warning does not come close to conveying this.

Pay attention to academic literature – as discussed in this article, an extensive academic literature shows that the SEC has too much faith in the capabilities of fund investors. Unfortunately, to date, the SEC has ignored this literature in regulating the industry. We examined five recent SEC rulemaking releases on mutual fund disclosure, which contain explanations and rationales for the rule proposals. In these five releases, the SEC cites a total of 339 sources in its discussion of investor behavior or information needs. The most frequently cited sources were comment letters from various individuals and organizations, including the ICI (75%), followed by news stories (8%), ICI research (4%), industry members (3%), SEC research (3%), non-academic research (2%), and ICI comments (2%). The SEC never cited academic research.

Thus, the SEC does not cite (much less rely upon) independent, academic studies indicating investors' needs and limitations, and whether investors can understand and will use the information that the SEC requires fund companies to disclose. Instead, the SEC has relied significantly on the fund industry for the agency's understanding of investor informational needs. For example, in its 1998 prospectus disclosure rulemaking, the SEC referred to ICI research reports seven times and ICI comment letters three times to support its conclusions about the information needs of fund investors.

The value of paying attention to academic research is not lost on everyone in the agency. In a recent speech, Brian Cartwright (General Counsel of the SEC) pointed out that "[g]iven the practical applications of theoretical finance, it's not surprising that today hedge fund managers regularly turn up at academic conferences and eagerly scour the academic literature for the next big thing. Regulators should too. If the investment choice most important for retail investors these days is not which stock or bond to buy, but which fund (or other intermediary) to choose, we need to find ways to apply the insights recent finance theory has given us in the service of retail investors." This would be a step in the right direction.

Conclusion

Retirement and saving for retirement are relatively recent phenomena. Seven decades ago our national retirement program was entrusted to a federal bureaucracy. Two decades later, our private retirement savings moved to company and government pension plans, administered by professional money managers. Today, as Social Security becomes less certain and defined-benefit pension plans give way to defined-contribution plans, our retirement savings has moved to mutual funds. In the process, fund investors have found themselves responsible for making essential investment choices concerning their retirement savings.

Are investors up to this heady new role? The fund industry, while painting a favorable portrait of fund investors, benefits from investor ignorance. The industry earns fees based on assets under management, and investors who pay little attention to fund costs create little pressure on fund companies to lower these costs. In addition, investors' tendency to chase past returns makes promoting past winners an effective marketing method. So long as investors pay insufficient attention to fund costs, the industry has an unassailable and highly profitable business model.

The SEC has undertaken only marginal reforms, focused primarily on streamlined disclosure and formulaic warnings. Relying significantly on the industry for information about investor needs, the SEC has largely accepted the story of a well-functioning market. While SEC reforms dance around the edges of investors' limitations, the agency seems in denial about the extent of this problem. Given the SEC's almost complete disregard of academic studies of fund investor sophistication and behavior, investors' needs are likely to be greatly underestimated and further reforms likely to be insufficient. As mutual fund investors, our financial future has been placed in our hands. And, as Pogo observed and the SEC has yet to fully recognize, "we have met the enemy and he is us."

Mutual fund investors: sharp enough?

References

- Alexander, G., J. Jones, and P. Nigro, 1998, "Mutual fund shareholders: characteristics, investor knowledge, and sources of information," *Financial Services Review*, 7, 301-316
- Barber, B., T. Odean, and L. Zheng, 2005, "Out of sight, out of mind: the effects of expenses on mutual fund flows," *Journal of Business*, 78, 2095-2120
- Barras, L., O. Scaillet, and R. Wermers, 2008, "False discoveries in mutual fund performance: measuring luck in estimated alphas, Working Paper, Robert H. Smith School
- Bergstresser, D., J. Chalmers, and P. Tufano, 2007, Assessing the costs and benefits of brokers in the mutual fund industry, Working Paper, Harvard Business School
- Berk, J., and R. Green, 2004, "Mutual fund flows and performance in rational markets," *Journal of Political Economy*, 112, 1269-1295
- Bogle, J., 2007, "Reflections on 'Toward common sense and common ground?'," *Journal of Corporate Law* 33, 31-40
- Capon, N., G. Fitzsimons, and R. Prince, 1996, "An individual level analysis of the mutual fund investment decision," *Journal of Financial Services Research*, 10, 59-82
- Choi, J., D. Laibson, and B. Madrian, 2006, "Why does the law of one price fail? An experiment on index mutual funds, working paper, Yale School of Management
- Coates, J., and R. Hubbard, 2007, "Competition in the mutual fund industry: evidence and implications for policy," *Journal of Corporate Law*, 33, 151-222
- Collins, S., 2005, "Are S&P 500 index mutual funds commodities?," *Investment Company Institute Perspectives* 11, 5
- Cooper, M., H. Gulen, and P. Rau, 2005, "Changing names with style: mutual fund name changes and their effects on fund flows," *Journal of Finance*, 60, 2825-2877
- Cox, J., and J. Payne, 2005, "Mutual fund expense disclosures: a behavioral perspective," *Washington University Law Quarterly*, 83, 907-938
- Cuthbertson, K., D. Nitzsche, and N. O'Sullivan, 2006, "Mutual fund performance," Working Paper, City University London
- Del Guercio, D., and P. Tkac, 2002, "The determinants of the flow of funds of managed portfolios: mutual funds vs. pension funds," *Journal of Finance & Quantitative Analysis*, 37, 523-557
- Elton, E., M. Gruber, and J. Busse, 2004, "Are investors rational? Choices among index funds," *Journal of Finance*, 59, 261-288
- Gallaher, S., R. Kaniel, and L. Starks, 2006, "Madison Avenue meets Wall Street: mutual fund families competition and advertising, Working paper, University of Texas at Austin
- Hortacsu, A., and C. Syverson, 2004, "Product differentiation, search costs, and competition in the mutual fund industry: a case study of S&P 500 index funds," *Quarterly Journal of Economics*, 119, 403-456.
- Huhmann, B., and N. Bhattacharya, 2005, "Does mutual fund advertising provide necessary investment information?," *International Journal of Bank Marketing*, 23, 296-313
- Investment Company Institute, 2008 "Investment Company Fact Book," [Fact Book]
- Investment Company Institute, 2008, "Profile of mutual fund shareholders," Spring, [Fund Shareholders]
- Investment Company Institute, 2008, "Research fundamentals: the U.S. retirement market, third quarter 2007," April, [Research Fundamentals]
- Investment Company Institute, 2008, "Trends in mutual fund investing, June, [Investing Trends]
- Investment Company Institute, 2007, "ICI research fundamentals: trends in ownership of mutual funds in the United States," November [Ownership Trends]
- Investment Company Institute, 2006, "Fees and expenses of mutual funds, 2005," [Fees & Expenses]
- Investment Company Institute, 2006, "Research commentary: competition in the mutual fund business, January, [Competition].
- Investment Company Institute, 2006, "Understanding investor preferences for mutual fund information," August, [Investor Preferences]
- Jain, P., and J. Wu, 2000, "Truth in mutual fund advertising: evidence on future performance and fund flows," *Journal of Finance*, 55, 937-958
- Jones, M., V. Llesseig, and T. Smythe, 2005, "Financial advisors and mutual fund selection," *Journal of Financial Planning*, 18, 64-70
- Khorana, A., and H. Servaes, 2004, "Conflicts of interest and competition in the mutual fund industry," Working Paper, Georgia Institute of Technology
- Mullainathan, S., and A. Shleifer, 2005, "Persuasion in finance," Working Paper, Harvard University
- Sirri, E., and P. Tufano, 1998, "Costly search and mutual fund flows," *Journal of Finance*, 53, 1589-1622
- Swensen, D., 2005, *Unconventional success: a fundamental approach to personal investment*, Free Press, New York, NY
- U.S. General Accounting Office, 2000, "Mutual fund fees: additional disclosure could encourage price competition," June, [GAO Mutual Fund Fees]
- U.S. Securities and Exchange Commission, 2000, "Report on mutual fund fees and expenses," December
- Wallison, P., and R. Litan, 2007, *Competitive equity: a better way to organize mutual funds*, American Enterprise Institute for Public Policy Research, Washington, D.C.
- Wilcox, R., 2003, "Bargain hunting or star gazing? Investors' preferences for stock mutual funds," *Journal of Business*, 76, 645-664
- Zhang, A., 2007, "Mutual fund expense ratios in market equilibrium," working paper, University of Arizona

82

MEMORANDUM

Date: September 29, 2014
To: Keith Bergstresser, U.S. Department of Labor
Cc: Youngok Lim (DOL), Michael Brien (Deloitte), Dirk Antonczyk (Deloitte)
From: Constantijn Panis (AACG)
Subject: Review of Study by Quantria Strategies, LLC*

SUMMARY

The Department of Labor (DOL) is expected to re-propose regulation on the definition of a fiduciary of employee benefit plans. Quantria Strategies, LLC (Quantria) conducted a study that attempts to estimate the effects of expanded fiduciary responsibilities on cash-outs of defined contribution (DC) plans by plan participants who separate from their job. Quantria estimated large increases in cash-outs and large reductions in the lifetime retirement savings of affected workers.

This document reviews the Quantria study. We point out that Quantria's estimates implicitly assumed that upcoming regulations will eradicate all financial advice on what to do with DC plan balances upon job separation, including by paid independent financial planners. We also note that Quantria relied on a correlation between financial advice and retirement assets to assert causality, where the causality may in fact go in the reverse direction. We further identify a key statistic about the average size of lump sum distributions that Quantria misrepresented as the average size of cash-outs and used to calculate long-term effects of cash-outs on lifetime retirement savings. Finally, we note that Quantria's conclusions on increased cash-outs and reduced lifetime retirement resources are based on apparently unrelated analyses, while the latter analysis applied assumptions that lack basis in empirical patterns.

* This document is the Final Analysis Memorandum, Deliverable 3.3 of Task Order DOL-OPS-14-T-00025 (Survey of Consumer Finances Statistical Studies and Literature Analysis) under Contract DOLJ139335157.

CONTENTS

Summary	1
1. Introduction	3
2. Synopsis of the Quantria Study	3
3. Discussion	5
Quantria Assumed That Re-Proposed Regulations Will Eradicate All Forms of Financial Advice on What to Do with Plan Balances upon Job Separation.....	5
Quantria Interpreted a Correlation as a Causal Effect	6
Quantria Misrepresented Lump Sum Distributions as Cash-Outs.....	7
There Is No Apparent Connection between Quantria’s Calculations of Incremental Cash-Outs and Reduced Lifetime Retirement Savings	8
4. Conclusion	9
Disclaimer	11

1. INTRODUCTION

The Employee Retirement Income Security Act of 1974 (ERISA) states that a person paid to provide investment advice with respect to assets of a private-sector employee benefit plan is a plan fiduciary. According to ERISA, a fiduciary must act solely in the interest of plan participants and their beneficiaries and with the exclusive purpose of providing benefits to them. Regulations promulgated by the U.S. Department of Labor (DOL) define who is a fiduciary. In 2010 DOL's Employee Benefits Security Administration (EBSA) proposed a new regulation that would expand the categories of persons who would be deemed to be fiduciaries subject to ERISA.¹ After receiving public comments, EBSA decided to re-propose its rule on the definition of a fiduciary.² As of September 2014, EBSA has not yet re-proposed such new regulation.

This document reviews a study by Quantria Strategies, LLC (Quantria) on "Access to Call Centers and Broker Dealers and Their Effects on Retirement Savings."³ The study asserts that the anticipated regulations will prompt financial service providers to limit the access to assistance provided by call centers and broker-dealers when terminating employees face plan distribution decisions. It finds that this could increase annual cash outs of retirement savings for employees terminating employment by \$20-32 billion and that these withdrawals could reduce the accumulated retirement savings of affected employees by 20%-40%.

The objective of this document is to review the soundness of the Quantria analysis, the quality of the data sources on which it is based, and the conclusions that it draws. We start in Section 2 with a summary of the Quantria study, its key assertions, logical steps, and conclusions. Section 3 addresses the study's arguments in light of findings in the academic and trade literature. Section 4 concludes.

2. SYNOPSIS OF THE QUANTRIA STUDY

The Quantria study explains that employees who participate in a defined contribution (DC) plan and separate from their job may have several options regarding the disposition of their DC plan balance. They can (a) preserve the assets for retirement by leaving them in their former employer's DC plan or rolling them over into an Individual Retirement Account (IRA) or (b) cash out their balance and spend the proceeds on anything they wish. Cashing out represents potential "leakage" of retirement resources because the assets may no longer be available for consumption in retirement.

The study warns that the re-proposed regulations that Quantria is anticipating from DOL may increase pension cash-outs and erode the retirement security of future retirees. Its conclusions are based on a number of key assertions and linkages—see Figure 1:

¹ Federal Register 75(204), pp. 65263-65278.

² <http://www.dol.gov/opa/media/press/ebsa/EBSA20111382.htm>.

³ The study is dated April 9, 2014 and was accessed through http://quantria.com/DistributionStudy_Quantria_4-1-14_final_pm.pdf.

1. DOL's upcoming regulations are expected to create fiduciary responsibilities for financial service providers. Concerned about potential liability, this may prompt financial service providers to limit access to assistance provided by call centers and broker-dealers when terminating employees face plan distribution decisions.
2. Call centers and broker-dealers may dissuade departing employees from cashing out their retirement account.
3. Cash-outs under the current system are substantial and may grow larger under the anticipated regulations, especially among individuals who have a low account balance, are under age 30, have lower wages, or are African-American or Hispanic.

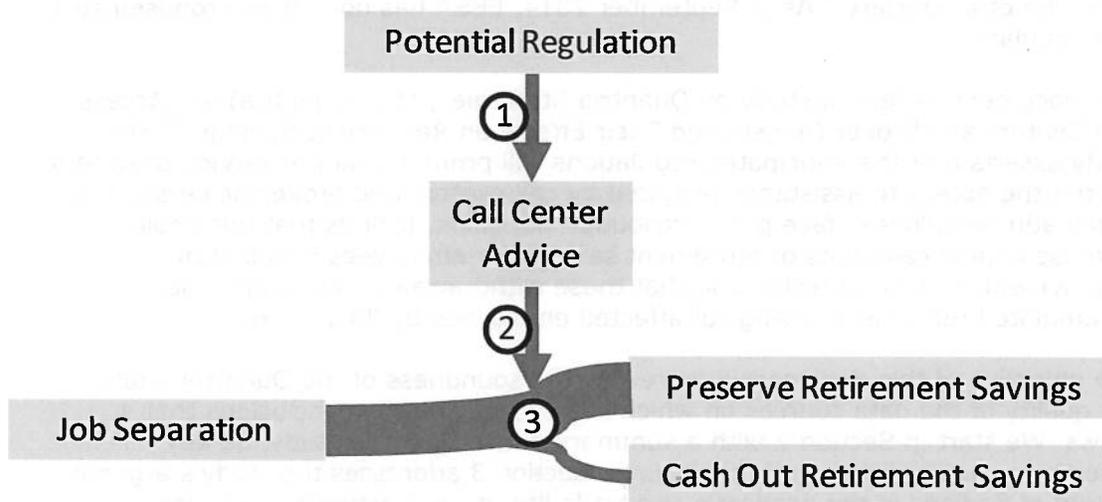


Figure 1. Causal Linkages Discussed in the Quantria Study

The Quantria study's primary conclusion is that anticipated regulations will increase annual cash outs of retirement savings at job termination by \$20-32 billion. The derivation of this result is not made explicit in the study, but pages 17-19 sketch the steps that Quantria analysts took:

1. Estimate pension participation by income category based on the number of taxpayers who deferred income and the aggregate amount of the deferral.
2. Impute aggregate balances for the distribution of taxpayers reporting retirement contributions on Form W-2.
3. Estimate the qualified plan participants who may experience a break in service (through job termination) and face a decision regarding plan assets held with their previous employer.
4. Derive estimates of the potential retirement savings balances that are likely to become vulnerable to cash out. Assume that the balances most vulnerable to increased cash-out are those currently being rolled into IRAs or a new employer plan and those that currently remain in an employer's plan through an employee decision. Quantria estimates that these balances total approximately \$117 to \$145 billion per year.
5. Estimate the effect of call center and broker-dealer assistance in helping workers retain their retirement savings at job change. Quantria estimates this effect at 33%.

In neither the main text nor its technical appendix does the Quantria study provide sufficient detail to convert at-risk account balances of \$117-145 billion per year and an assistance effect of 33% into its main conclusion. The study merely states the following on page 19:

“Based on the estimated cash-outs from retirement balances for terminating employees, we estimate that reducing the availability of call center and broker-dealer assistance will increase annual cash outs of retirement savings at job termination by an additional \$20-32 billion.”

However, Quantria notes that plan balances increase with age initially and then increase at a decreasing rate as workers approach retirement age. It may be the case that Quantria applied its estimated assistance effect to age-specific account balances to derive its aggregate effect of \$20-32 billion.

Quantria also simulates the effect of cash-outs on lifetime retirement savings for two scenarios. In the first the authors assume that a worker cashes out all plan balances during the first six years of his career and find that that could reduce lifetime retirement savings by 24%. In the second the authors assume that a worker cashes out a plan balance approximately equal to the average lump sum distribution in 2006 and find that that could reduce lifetime retirement savings by 41%. Based on these simulations Quantria concludes that the incremental cash-outs of \$20-32 billion could reduce the ultimate retirement savings of affected individuals by 20% to 40%.

3. DISCUSSION

The Quantria study relies on statistics from several external research papers and several data sets for its own analysis. It combines those statistics and data sets even though these components are not always compatible or consistent. It also misinterprets several statistics and relationships. Below we highlight the most obvious errors.

Quantria Assumed That Re-Proposed Regulations Will Eradicate All Forms of Financial Advice on What to Do with Plan Balances upon Job Separation

The Quantria study starts off with the following statement (page 1):

Re-proposed regulations that are anticipated from the Department of Labor (DOL) are generally expected to create fiduciary responsibility for financial service providers that will limit the access to assistance provided by call centers and broker-dealers when terminating employees face plan distribution decisions.

The study appears to be based on this central premise and does not investigate whether this premise is plausible, to what extent upcoming regulation will limit access to assistance provided by call centers and broker-dealers, or to what extent plan participants will find substitutes for any lost advice.

Investment firms aim to profit from the asset holdings of their clients and thus have an incentive to dissuade a separating employee from cashing out his account balance. For example, Sholder (2012) estimated that retaining a plan participant

after retirement could result in 32%-53% of the participant's lifetime value to a financial service provider.⁴ By extension, retaining a participant upon pre-retirement job separation could result in an even greater share of his lifetime value. It may therefore be expected that financial service providers will continue to court departing plan participants under new fiduciary regulations.

To derive its result of \$20-32 billion higher cash-outs under expected regulations, Quantria relied on its "theory that consultations with call centers or brokers/dealers result in retirement savings in DC plans that are higher by about 33 percent" (page 19). That theory is apparently based on a model showing that retirement savings plan balances in a national household survey were 33 percent higher for respondents who had consulted a financial planner or broker for financial advice than for respondents who had not. Quantria acknowledges in a footnote that paid financial planners could have been consulted by survey respondents, rather than only call center operators or broker-dealers (footnote 43, page 18). However, its analysis appears to have applied the 33% estimate in full and thus ignored that some of the correlation arose from consultations with financial planners. Implicitly, Quantria assumed that upcoming regulation will eradicate all financial advice on what to do with plan balances upon job separation from financial planners and brokers. Its assumption goes beyond advice from call centers and broker-dealers only. To the extent that the upcoming regulation will reduce financial advice, Quantria even assumed that plan participants will not find substitutes for such lost advice.

Quantria Interpreted a Correlation as a Causal Effect

Quantria estimated a statistical model of retirement account balances based on data from the RAND American Life Panel (ALP). It found that respondents who reported having consulted financial planners or brokers owned accounts with balances that, on average, were 33% higher than respondents who had not consulted a financial planner or broker. This is a correlation, not necessarily a causal effect. Quantria first reported this estimate as a correlation but then its language evolved into a causal interpretation:

- Page 18-19: "Our empirical models suggest that retirement savings plan balances are 33 percent higher if a financial planner or broker was consulted for financial advice."
- Page 19: "Our parameter estimates support our theory that consultations with call centers or brokers/dealers result in retirement savings in DC plans that are higher by about 33 percent."
- Page 38: "In particular, having access to a financial planner or broker appears to result in retirement savings in DC plans that are higher by about 33 percent."

As pointed out by Munnell (2014), the model does not support the assertion that consulting a financial planner "results" in more wealth, and hence the loss of access

⁴ James R. Sholder (2012): "Retaining Retirement Plan Rollovers: Rationale and Process." The Diversified Services Group, Inc. Retrieved from http://www.dsg-network.com/pdf/DSG-White-Paper-Retaining-Retirement-Plan-Rollovers-March2012_fx.pdf.

to the financial planner would cause a reduction in wealth.⁵ The causation may run in the other direction—individuals who are thrifty, wealthy, or financially sophisticated to start with may find it more worthwhile to hire a financial planner. If they lost their financial planner, they might invest a little less successfully, but they would still hold relatively high balances.

Separately, Quantria assumed that whatever caused account balances to be 33% higher translates into an equal-sized effect on preservation of disposable funds upon job separation. In other words, it applied a stock concept (balances) to a flow (annual rollovers and other dispositions).

Quantria Misrepresented Lump Sum Distributions as Cash-Outs

The Quantria study cited a study by Aon Hewitt showing that 42% of employees take a cash distribution of their retirement savings at job termination, 29% roll their retirement savings to another plan or an IRA, and 29% leave their assets in the employer's plan (page 11).⁶ It proceeded to further quantify cash-outs with a metric from a study of the Employee Benefit Research Institute (EBRI)⁷ (page 12):

"Cashing out tends to be the most prevalent behavior with average distributions of \$32,219 (in 2006 dollars)."

The EBRI (2009) study found that the average lump sum distribution was \$32,219 in 2006. However, a lump sum distribution is not the same as a cash-out. A lump sum distribution is a one-time payment which may be used to roll over the funds into an IRA or be cashed out.⁸ The average \$32,219 lump sum distribution is thus a mix of rollovers and cash-outs.

As is widely documented in the literature, average rollovers tend to be much greater than average cash-outs. For example, the same Aon-Hewitt study that was cited by Quantria found that 42% of participants cashed-out, but those cash-outs represented only 7.3% of assets. By contrast, 29% of participants rolled their balance over into an IRA, and those rollovers represented 38% of assets. More recently, Vanguard, a financial services provider, reported that cash-outs by 28% of participants represented only 5% of assets in 2013, whereas rollovers by 22% of participants accounted for 39% of assets.⁹

⁵ Alicia H. Munnell (2014): "Fiduciary rules would help, not hurt, savers." Retrieved from <http://blogs.marketwatch.com/encore/2014/05/07/fiduciary-rules-would-help-not-hurt-investors>.

⁶ Aon Hewitt (2011): "Leakage of Participants' DC Assets: How Loans, Withdrawals, and Cashouts Are Eroding Retirement Income 2011." Retrieved from http://www.aon.com/attachments/thought-leadership/survey_asset_leakage.pdf.

⁷ "Lump-Sum Distributions at Job Change," *EBRI Notes* Vol. 30, No. 1, January 2009. Available at http://www.ebri.org/pdf/notespdf/EBRI_Notes_Jan09_Rollovers.pdf.

⁸ It may also be used to purchase an annuity or to be rolled over into the plan of the participant's new employer.

⁹ "How America Saves, 2014," The Vanguard Group, Inc. Retrieved from <https://institutional.vanguard.com/iam/pdf/HAS14.pdf>.

Based on the participant and asset fractions reported in the Aon-Hewitt study, the average rollover was approximately 7.5 times as large as the average cash-out.¹⁰ If the average lump sum distribution was \$32,219, this would imply that the average cash-out was approximately \$8,777. However, Quantria selected an average lump sum distribution from the EBRI report that was based on historical data going back to before 1980; the most recent figure reported in the EBRI study was \$22,166 (for 2004-2006) and the trend was downward.

In short, EBRI reported a time series of average lump sum distributions. Quantria selected the historical average rather than the much-lower average over the most recent period and further misrepresented it as the average cash-out amount. As discussed in the next subsection, this implies that Quantria greatly inflated its long-term effects on lifetime retirement savings.

There Is No Apparent Connection between Quantria's Calculations of Incremental Cash-Outs and Reduced Lifetime Retirement Savings

The Quantria study states (page 4):

"We estimate that eliminating the availability of call centers and reducing broker-dealer assistance upon job termination will increase annual cash outs of retirement savings by an additional \$20-32 billion. Over the long run, these cash outs will result in a significant reduction in overall retirement savings; our estimates indicate that these withdrawals could reduce the ultimate retirement savings of affected individuals by 20 to 40 percent."

The language seems to suggest that the increase in cash-outs by \$20-32 billion corresponds to a reduction of ultimate retirement savings of 20%-40%. However, the two results appear to be based on unrelated calculations.

The preceding subsections explain how Quantria estimated increased cash-outs of \$20-32 billion. The reduction in lifetime retirement savings is based on simulations of hypothetical DC accumulation paths.

Quantria assumed a baseline scenario of a worker who earns \$40,000 in 2014. Her wage increases by 2.5% annually, she contributes 5% of her earnings to her DC plan, and she earns 4% interest on her plan balance. According to Quantria's calculations, this worker will accumulate \$124,742 during her career, which is assumed to end in 2040 (page 19).

For its first alternative scenario, Quantria simulated the accumulations of a similarly situated worker who cashes out all balances during the first six years of her career. Her end-of-career accumulation would be \$92,650, i.e., a reduction of 26%.

This first alternative scenario captures cash-outs early in someone's career. For its second alternative scenario, Quantria noted (page 20):

¹⁰ The ratio of 38%/29% and 42%/7.3% is $5.75/0.76 = 7.5$. Based on Vanguard's figures, the average rollover amount is even larger, namely almost ten times the size of the average cash-out amount.

While it is common for younger workers to cash out balances, the average account distribution is considerably higher (\$32,219) suggesting that cash outs also occur later in a working career. Using the same example, this is comparable to a worker cashing out his or her retirement savings when changing jobs after 10 years. The effect of this behavior is to reduce the individual's total retirement savings by nearly \$52,000 or 41 percent.

The second alternative scenario's design is based on the premise that the average cash-out is \$32,219. (As explained above, this figure is greatly inflated.)

These simulations are the only calculations in the Quantria study that aim to calculate long-term effects of cash-outs. In the Executive Summary and the Findings section, the results are summarized as a reduction in ultimate retirement savings of affected individuals by 20%-40%. Neither alternative scenario is based on or connected to an increase in cash-outs by \$20-32 billion. In fact, neither appears to be grounded in empirical patterns. The first alternative scenario simply assumes that accumulations during the first six years of someone's career are cashed out; the second is loosely based on a greatly inflated number, namely an obsolete average lump sum distribution amount that is misrepresented as an average cash-out amount.

Finally, Quantria assumed that cashed out funds are not available for consumption in retirement. It ignored the possibility that cashed out balances can be saved or used to start a business or purchase a home with benefits that may stretch into retirement. For example, EBRI (2009) reported that lump sum distributions were about twice as likely to be used for paying off debts, funding a business, or purchasing a home than for consumption.¹¹

4. CONCLUSION

The Quantria study estimated that expected DOL regulations concerning the definition of a fiduciary will trigger large increases in cash-outs of DC plan balances upon job separation and large reductions in lifetime retirement resources. Our review demonstrates that the study is based on flawed or arbitrary assumptions, flawed methods, and misrepresentations of external findings.

That said, the basic idea behind the study—that expanded fiduciary responsibilities may reduce financial advice and increase cash-outs—is possible, but just not empirically proven or justified from the Quantria report. Depending on the content of the regulation (which is still unknown), there may be a modest increase in cash-outs. One perspective is that the benefits of less expensive investment products for many American workers may outweigh the costs of higher cash-outs by some individuals. Another perspective, expressed by the General Accounting Office (GAO), holds that the benefits of less expensive investment products may be achieved while avoiding or limiting increased cash-outs. In a response to the Quantria study, the GAO expressed the hope that any regulation the DOL promulgates will clarify the distinction between investment advice that may trigger fiduciary status and

¹¹ Figures 5 and 6 of the EBRI (2009) study referenced in footnote 7.

investment education that will not, thereby mitigating any concern that call centers will no longer be able to provide helpful guidance to separating plan participants.¹²

Indeed, the central premise of the Quantria study will be removed if the regulation manages to keep in place or even improve guidance to separating plan participants.

¹² Letter of 28 July 2014 from Charles A. Jeszeck, Director of GAO's Education, Workforce, and Income Security Team to Congressman George Miller, Ranking Member of the Committee on Education and the Workforce in the House of Representatives.

DISCLAIMER

The views, opinions, and/or findings contained in this report are those of the authors and should not be construed as an official Government position, policy or decision, unless so designated by other documentation issued by the appropriate governmental authority.

We call your attention to the possibility that other professionals may perform procedures concerning the same information or data and reach different findings than Advanced Analytical Consulting Group, Inc. (AACG) and Deloitte Financial Advisory Services LLP (Deloitte) for a variety of reasons, including the possibilities that additional or different information or data might be provided to them that was not provided to AACG and Deloitte, that they might perform different procedures than did AACG and Deloitte, or that professional judgments concerning complex, unusual, or poorly documented matters may differ.

This document contains general information only. AACG and Deloitte are not, by means of this document, rendering business, financial, investment, or other professional advice or services. This document is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action. Before making any decision or taking any action, a qualified professional advisor should be consulted. AACG and Deloitte, its affiliates, or related entities shall not be responsible for any loss sustained by any person who relies on this publication.

83

Financial Asset Holdings of Households in the United States in 2013

January 21, 2016

Constantijn W.A. Panis, PhD

Advanced Analytical Consulting Group, Inc.

213-784-6400

stanpanis@aacg.com

Michael J. Brien, PhD

Deloitte Transactions and Business Analytics LLP

202-378-5096

michaelbrien@deloitte.com

ABSTRACT

This report aims to support the U.S. Department of Labor (DOL) with tabulations of asset holdings and investment patterns of households in the United States. It is organized in two parts. Part I contains a summary of publicly available data sources with measures of holdings of 401(k) and other defined contribution (DC) plans, individual retirement accounts (IRAs), and after-tax investment accounts. The primary focus is on the Survey of Consumer Finances and the Survey of Income and Program Participation. Part II contains tabulations and analyses based on the Survey of Consumer Finances of 2013. It documents differences among owners of IRAs, 401(k) and other DC plans, and after-tax investment accounts.

CONTENTS

Abstract	i
Part I. Summary of Available Data	1
I.1. Introduction	2
I.2. Survey of Consumer Finances	3
Overview	3
Survey Design and Weights	4
Imputation	4
Survey Content	5
Summary Statistics	7
I.3. Survey of Income and Program Participation	10
Overview	10
Survey Design and Weights	10
Topical Modules	11
Summary Statistics	14
I.4. Other Data Sources	16
Health and Retirement Study	16
Panel Study of Income Dynamics	16
Employee Benefit Research Institute Data	17
The Center for Research in Security Prices Data	17
Financial Research Corporation Data	18
E-filings of SEC Form N-SAR	19
FINRA BrokerCheck Reports	19
SEC Form ADV	20
State Financial Regulators	22
I.5. Bibliography	23
Part II. Statistical Analysis	24
II.1. Introduction	25
II.2. Tabulations and Analyses	26
Appendix 1	60
Availability of Data on IRA and DC Account Age	60
Appendix 2	61
Calculations for Table 26: IRAs, IRA assets, and Median Household Income of IRA Holders, by Type of Institution Holding IRAs	61
Disclaimer	62

TABLES

Table 1: SCF Data Availability	4
Table 2: IRA/Keogh Accounts.....	5
Table 3: Pensions from Current Main Job	6
Table 4: Pensions from Past Jobs	6
Table 5. Asset Allocation of First DC Plan from Respondent's Current Job	7
Table 6. Asset Allocation of Respondent's First IRA/Keogh Account.....	7
Table 7. First Institution where Respondent's IRAs Are Held	8
Table 8. Sources of Investment/Savings Advice.....	8
Table 9. Demographic Characteristics.....	9
Table 10. Topical Modules Related to Retirement and Assets	11
Table 11. Topical Modules in the 2008 Panel	13
Table 12. Asset Types Reported in IRAs (2008 Panel, Wave 5).....	14
Table 13. Asset Types Reported in DC Plans (2008 Panel, Wave 5).....	14
Table 14. Demographic Characteristics (2008 Panel, Wave 3).....	15
Table 15. Categories of Households by Ownership and Value of Account	26
Table 16. Categories of Households by Ownership and Value of Accounts (including all retail investments and savings).....	27
Table 17. Taxable Asset Holders and Holdings by Asset Type.....	28
Table 18. Income Distribution for All Households and for Those with IRAs, DC Plans, or Taxable Accounts	29
Table 19. Income Distribution of Asset Holdings for Households with IRAs, DC Plans, or Taxable Accounts	31
Table 20. Income Summary Statistics for Households with IRAs, DC Plans, or Taxable Accounts (\$ thousands).....	32
Table 21. Account-Balance Summary Statistics for IRAs, DC Plans, and Taxable Accounts (\$ thousands).....	32
Table 22. Stock-Fraction Summary Statistics for IRAs, DC Plans, and Taxable Accounts.....	32
Table 23. Distribution of Demographic Characteristics of the Heads of All Households and of Those with IRAs, DC Plans, or Taxable Accounts.....	33
Table 24. Asset-Weighted Distribution of Demographic Characteristics of the Heads of Households with IRAs, DC Plans, or Taxable Accounts	34
Table 25. Margin Loan Statistics.....	34
Table 26: IRAs, IRA assets, and Median Household Income of IRA Holders, by Type of Institution Holding IRAs.....	35
Table 27. Distribution of Asset Holdings By Household Income.....	36
Table 28. Distribution of Households, IRA Accounts, and IRA Assets by Total Household IRA Assets	39
Table 29. Distribution of Number of IRA Accounts per Household Member by Total Household IRA Assets	40
Table 30. Household Account Types by Source of Financial Advice.....	41

Table 31. Sources of Financial Advice by Household Account Types	42
Table 32. Sources of Financial Advice for Households with an IRA, by Household IRA Assets	43
Table 33. Sources of Financial Advice for Households with an IRA, by Household IRA Assets	44
Table 34. IRA Assets of Households That Consulted a Financial Planner, Banker, Broker, Accountant, or Lawyer	45
Table 35. IRA Assets of Households That Consulted a Broker	46
Table 36. Median Household Income, by IRA Ownership and Professional Advice....	47
Table 37. Sources of Financial Advice by Household Income	48
Table 38. Sources of Financial Advice by Head of Household Demographics	49
Table 39. Sources of Financial Advice by Financial Institution(s) that Hold Households' IRAs	50
Table 40. Households' IRA-holding Financial Institution(s) by Total Household IRA Assets	51
Table 41. Households' IRA-holding Financial Institution(s) by Total Household Income	52
Table 42. Households' IRA-holding Financial Institution(s) by Head of Household Demographics.....	53
Table 43. Households' IRA-holding Financial Institution(s) by Total Household Wealth Quintiles and Top Percentiles	54
Table 44. Distribution of Asset Holdings by Total Household Wealth Quintiles.....	55
Table 45. Distribution of Asset Holdings Overall and by Total Household Wealth Top Percentiles	57
Table 46. Distribution of Asset Holdings among IRA Holders by Total IRA-holding Household Wealth Quintiles	58
Table 47. Distribution of Asset Holdings among IRA Holders Overall and by Total IRA-holding Household Wealth Top Percentiles.....	59

FIGURES

Figure 1. Income Distribution for All Households and for Those with IRAs, DC Plans, or Taxable Accounts	30
Figure 2. Household Income Distribution of Asset Holdings for Households with IRAs, DC Plans, or Taxable Accounts	31
Figure 3. Household Asset Holdings, Income < \$10K.....	37
Figure 4. Household Asset Holdings, Income \$10K-20K	37
Figure 5. Household Asset Holdings, Income \$20K-50K	37
Figure 6. Household Asset Holdings, Income \$50K-100K.....	37
Figure 7. Household Asset Holdings, Income \$100K-200K	38
Figure 8. Household Asset Holdings, Income \$200K-500K	38
Figure 9. Household Asset Holdings, Income \$500K+	38
Figure 10. Household Asset Holdings, Total Population	38

PART I. SUMMARY OF AVAILABLE DATA

I.1. INTRODUCTION

Part I of this report summarizes data that are available to measure holdings of 401(k) and other defined contribution (DC) plans, individual retirement accounts (IRAs), and after-tax investment accounts. The emphasis is on data sources which may be particularly useful in supporting an analysis of the potential reach of proposed changes in the definition of a fiduciary.

The main data sources of interest are the Survey of Consumer Finances (SCF) and the Survey of Income and Program Participation (SIPP). The SCF is a repeated cross-section of household-level data, collected every three years. It combines a geographically and socio-economically representative sample with an oversample of wealthy households. See Section I.2 for a description of the SCF, including its survey design, sample weighting scheme, availability, and key questions about investments. The SIPP is a repeated panel survey. About every three years, a new multi-stage stratified sample is drawn, and respondents are interviewed approximately every four months for about four years. Section I.3 discusses the SIPP, including its survey design, core and modular topics, availability, and key questions about investments.

Section I.4 discusses other datasets that are potentially of interest. The Health and Retirement Study (HRS) is a panel survey of Americans over age 50 interviewed every other year. Every six years, the sample is augmented with respondents from younger cohorts. The Panel Study of Income Dynamics (PSID) is a panel survey which dates back to 1968. Respondents were interviewed annually through 1997 and every other year since 1999. The PSID attempts to follow not just original respondents but also others who split off from the respondent's household. Employee Benefit Research Institute (EBRI) data are compiled from surveys of IRA plan administrators as well as of the Investment Company Institute. Center for Research in Security Prices (CRSP) data give daily stock prices, dividends and rates of return since 1926. Financial Research Corporation (FRC) data and Securities and Exchange Commission (SEC) filings of Semi-Annual Reports for Registered Investment Companies (Form N-SAR) are also briefly discussed.

I.2. SURVEY OF CONSUMER FINANCES

Website: <http://www.federalreserve.gov/pubs/oss/oss2/about.html>

Overview

The SCF is conducted every three years and gathers information on various financial and demographic characteristics of families in the United States. The study is sponsored by the Federal Reserve Board (FRB) in cooperation with the Department of the Treasury, and data are collected by the National Organization for Research at the University of Chicago (NORC).

SCF data are intended to represent the "primary economic unit" (PEU). The PEU is defined as a single individual or couple (married or living as partners) and all individuals who are financially dependent on that individual or couple.¹ In most instances, the PEU and the household are identical.

The last three survey years for which data are currently publicly available are 2007, 2010, and 2013. About 4,500 households were sampled in 2007, 6,500 in 2010, and 6,000 in 2013. The SCF includes some respondents who qualify for the Forbes Magazine list of the 400 richest Americans, but these respondents are excluded from the public data set.

In order to examine the effect of the recent financial crisis and ensuing recession, a special SCF was fielded in 2009 that attempted to re-interview respondents to the 2007 SCF.² The 2009 follow-up SCF included fewer questions than the 2007 survey and focused on topic areas most helpful for understanding the financial crisis' impact on families in the United States (Bricker et al., 2011; e-mail communication with FRB staff). Specifically, the highest-level questions on different components of wealth were retained; but, in most instances, fewer detailed questions were asked.

Data for the 1983-2013 waves of the SCF along with the 2007-2009 panel survey are publicly available online. Table 1 summarizes SCF data availability beginning with the 2001 survey.

¹ For example, in a household that consists of a married couple, two dependent children, and the financially independent parent of one of the spouses, the PEU would consist of only the married couple and two children.

Table 1: SCF Data Availability

Survey year	Date available	Number of respondents (publicly available)
2001	currently available	4,442
2004	currently available	4,519
2007	currently available	4,418
2007-2009 Panel	currently available	3,857
2010	currently available	6,482
2013	currently available	6,015

Sources: <http://www.federalreserve.gov/econresdata/scf/scfindex.htm>
 SCF Codebook (2001, 2004, 2007, 2007-2009 Panel, 2010, 2013)

The SCF utilizes computer-assisted personal interviewing (CAPI). Interviews are conducted in person or over the phone and in English or in Spanish. In 2013, about 34 percent of the interviews were conducted over the phone, and the vast majority were conducted in English.

Survey Design and Weights

The SCF employs a dual-frame sample design that is made up of a national multi-stage area probability (AP) sample and a list sample that oversamples wealthy households. The purpose of the dual-frame design is to ensure that information is collected both on attributes that are broadly distributed throughout the population and on those that are narrowly focused towards the wealthy. In 2013, approximately four-fifths of respondents originate from the AP sample and one-fifth from the list sample.

The AP sample focuses on characteristics that are broadly distributed in the population, such as credit card debt. The sample is selected in several stages of clustering and stratification. First the U.S. is divided into geographic groups. Then some areas are selected as Primary Sampling Units (PSUs) with probability one, while others are drawn randomly from geographically based strata. Further stratification defines sub-areas within PSUs, and from the smallest of these (roughly "block" level) households are sampled.

The list sample is selected from the Individual Tax File (ITF), an annual probability sample drawn from the complete set of tax returns by the Statistics of Income Division of the Internal Revenue Service. The list sample oversamples wealthy households, boosting sample coverage of variables that are largely concentrated among wealthy households, such as ownership of corporate stock.

The dual-frame sampling procedure, along with several instances of clustering and stratifying, require analysis weights to be computed for the SCF data. The SCF computes a single analysis weight for the complete sample of households. For details see Kennickell and Woodburn (1997) or Kennickell (1999, 2000).

Imputation

Most missing data items have values imputed. The imputation process does not correct for certain hierarchical inconsistencies. For example, total income in the

observed data is sometimes not equal to the sum of the individual components, and imputed data may likewise exhibit such an inconsistency.

The imputations are stored as five replicates of the data, based on five different imputation assumptions. In other words, the number of observations in the public file is five times as large as the number of respondents, and non-missing data have equal values (that is, the observed value) in all five replicates. Data items that are imputed are flagged as such. A detailed description of the algorithm used for imputations can be found in Kennickell (1998).

Survey Content

The SCF asks questions on DC pensions and IRAs separately. The following tables list the paraphrasing of questions that may be of interest for analysis, and their availability in the 2007, 2010 and 2013 surveys.

Table 2: IRA/Keogh Accounts

Question	Survey Year		
	2007	2010	2013
How many IRA and Keogh accounts do you have? (up to 5 recorded)	x	x	x
Are these accounts:	x	x	x
• Roth IRAs			
• rolled over from other pension accounts?			
• regular IRAs?			
• Keoghs?			
How much is in your:	x	x	x
• Roth IRAs			
• rolled over from other pension accounts?			
• regular IRAs?			
• Keoghs?			
How is the money in this account invested?	x	x	x
• stocks			
• interest-earning assets			
• split			
• something else			
Where are these accounts held?	x	x	x
• commercial bank			
• savings and loan/savings bank			
• credit union			
• brokerage			
• other			

Table 3: Pensions from Current Main Job

Question	Survey Year		
	2007	2010	2013
Are you included in any pension, retirement, or tax-deferred savings plans connected with your current main job? (detailed information on up to 6 plans, summary information on all other plans)	x	x	x
Is the plan the type in which money accumulates in an account designated for the worker and the worker or employer may make contributions to the account?	x	x	x
Or, is the plan a traditional pension plan that provides regular benefits at retirement based on years of work and pay?	x	x	x
Do you have a choice about how these funds are invested?	x	x	x
How is the money in this account invested?	x	x	x
<ul style="list-style-type: none"> • stocks • interest-earning assets • split • something else 			
What percent of this account is invested in stocks?	x	x	x
Is any of this stock in the company where you work?	x	x	x
What percent of your pay do you contribute to the plan?	x	x	x
What percent of your pay does your employer contribute to the plan?	x	x	x

Table 4: Pensions from Past Jobs

Question	Survey Year		
	2007	2010	2013
Do you have any accounts from a previous employer that you will receive or draw on in the future? (detailed information on up to 6 accounts, summary information on all others)	x	x	x
About this benefit, is it a type of plan that has an account balance, is it a plan that will pay out a regular income for life after retirement, or is it a mixture of the two types?	x	x	x
How much is in the account now?	x	x	x
How is the money in this account invested?	x	x	x
<ul style="list-style-type: none"> • stocks • interest-earning assets • split • something else 			
What percent of this account is invested in stocks?	x	x	x
Is any invested in the stock of the company where you worked?	x	x	x

Independent from any type of account, the SCF further inquires about the sources of advice on savings or investments which respondents use.

Summary Statistics

This section presents illustrative summary statistics on demographics, asset allocation, institutions in which IRAs are held, and investment advice. All tabulations are based on the 2013 SCF.

Table 5 tabulates the asset allocation of the respondent's first DC plan. Similarly, Table 6 tabulates the asset allocation of the respondent's first IRA or Keogh account.

Table 5. Asset Allocation of First DC Plan from Respondent's Current Job

	Freq.	Percent	Weighted Percent
Other	5	0.1%	0.0%
All in stocks	2,245	26.4%	23.6%
All in interest earning assets	962	11.3%	12.0%
Split	5,303	62.3%	64.4%
Total	8,515	100.0%	100.0%

Source: 2013 SCF.

Table 6. Asset Allocation of Respondent's First IRA/Keogh Account

	Freq.	Percent	Weighted Percent
Other	75	0.8%	0.7%
All in stocks	3,284	34.3%	33.0%
All in interest earning assets	1,210	12.6%	14.9%
Split	5,019	52.3%	51.4%
Total	9,588	100.0%	100.0%

Source: 2013 SCF.

Table 7 tabulates the first institution mentioned by the respondent.

Table 7. First Institution where Respondent's IRAs Are Held

	Freq.	Percent	Weighted Percent
Commercial Bank	4,377	45.7%	45.2%
Savings Loan/Bank	198	2.1%	3.0%
Credit Union	432	4.5%	7.2%
Finance/Loan Company	132	1.4%	1.4%
Brokerage	4,143	43.2%	39.9%
Insurance Company	143	1.5%	1.6%
Mortgage Company	25	0.3%	0.3%
Accountant	20	0.2%	0.2%
Current or Former Employer	5	0.1%	0.1%
Individual Formal Lender	10	0.1%	0.1%
Pension Administrator	10	0.1%	0.1%
Internet-based Business	5	0.1%	0.0%
Self/Spouse/Partner	6	0.1%	0.1%
Fraternal Organization	5	0.1%	0.1%
Other Member Org	5	0.1%	0.1%
Investment/Management Company	72	0.8%	0.7%
Total	9,588	100.0%	100.0%

Source: 2013 SCF.

Table 8 tabulates what sources of information respondents (and their family) use to make decisions about saving and investments.

Table 8. Sources of Investment/Savings Advice

Source	Freq.	Percent	Weighted Percent
Call around	3,881	12.9%	13.2%
Magazines/Newspapers/Books	4,086	13.6%	11.0%
Material in the mail	1,809	6.0%	6.4%
Television/Radio	2,402	8.0%	8.0%
Internet/Online services	11,269	37.5%	35.3%
Advertisements	1,783	5.9%	6.2%
Friends/Relatives	11,565	38.5%	40.1%
Lawyers	1,574	5.2%	3.8%
Accountants	4,086	13.6%	10.2%
Bankers	10,082	33.5%	33.0%
Brokers	3,829	12.7%	8.6%
Financial planners	9,034	30.0%	25.5%
Self	2,438	8.1%	6.8%

Source: 2013 SCF.

Note: Percentages do not add to 100% because multiple answers were allowed.

Table 9 reports summary statistics on various demographic characteristics in the 2013 SCF.

Table 9. Demographic Characteristics

	Freq.	Percent	Weighted Percent
Marital status			
Married	16,430	54.6%	48.1%
Separated	1,005	3.3%	3.5%
Divorced	4,602	15.3%	16.6%
Widowed	2,230	7.4%	9.6%
Never married	5,808	19.3%	22.1%
Education			
High school or less	2,720	9.0%	11.0%
High school or equivalency	8,006	26.6%	31.3%
Some college	5,139	17.1%	18.9%
Associate's degree	1,847	6.1%	6.6%
Bachelor's degree	6,730	22.4%	19.7%
Master's degree	3,631	12.1%	9.3%
Professional degree/Doctorate	2,002	6.7%	3.2%
Ethnicity			
Hispanic/Latino origin	3,562	11.8%	13.3%
Race			
White	22,123	73.6%	70.1%
Black	3,731	12.4%	14.6%
Hispanic	2,781	9.2%	10.6%
Other	1,440	4.8%	4.7%
Sex			
Male	22,960	76.3%	71.6%
Female	7,115	23.7%	28.4%
Age			
18-25	1,440	4.8%	6.2%
26-35	4,140	13.8%	16.2%
36-45	5,325	17.7%	17.5%
46-55	6,710	22.3%	19.7%
56-65	6,295	20.9%	18.6%
66-75	3,795	12.6%	12.2%
76-85	1,805	6.0%	7.5%
>85	565	1.9%	2.3%

Source: 2013 SCF.

I.3. SURVEY OF INCOME AND PROGRAM PARTICIPATION

Website: <http://www.census.gov/sipp>

Overview

Collected by the U.S. Census Bureau, the SIPP collects data on income, labor-force status, and welfare program participation and eligibility, along with general demographic characteristics. This facilitates measuring the effectiveness of existing federal, state, and local programs; estimating future costs and coverage for government programs, such as food stamps; and providing improved statistics on the distribution of income and on measures of economic well-being (U.S. Census Bureau, 2009). Low-income households are oversampled in the SIPP.

All of this report's SIPP-related facts, statements, and tabulations are current as of September 30, 2011.

Survey Design and Weights

The survey is a continuous series of national panels with a target sample size of 37,000 households. Each panel lasts from 2½ years to 4 years, during which respondents are interviewed approximately every four months. The most recent three panels started interviews in 2001, 2004, and 2008; the 2008 panel is currently active.

The SIPP is a multi-stage stratified sample of the U.S. civilian non-institutionalized population. It attempts to interview every household member age 15 or older. Core questions address labor force participation, program participation and income designed to measure a household's economic situation. These questions are repeated in every wave to monitor the distribution of cash and assets within the United States. Both telephone interviews and personal visits are used to complete the interview.

The SIPP also contains a series of topical modules with questions on a variety of targeted topics, including on Income & Retirement Accounts, Retirement & Pension Plans, and Assets & Liabilities. These questions are assigned to particular interview waves of the survey. In addition to retirement topics, topical modules cover personal history, child care, program eligibility, child support, utilization and cost of health care, disability, school enrollment, taxes, and annual income (U.S. Census Bureau, 2009; Westat et al., 2001).

Table 10 lists modules of particular interest, along with the number of respondents, in the 2001, 2004, and 2008 SIPP panels. Note that questions on various types of assets tend to be administered several times to every panel, at three-wave (one-year) intervals. Under suitable assumptions, this may permit calculations of rates of return on account assets.

Table 10. Topical Modules Related to Retirement and Assets

Panel	Wave	Interview period	Module	Respondents
2001	3	Oct 2001-Jan 2002	Assets, Liabilities	71,280
	4	Feb-May 2002	Income, Retirement Accts	70,595
	6	Oct 2002-Jan 2003	Assets, Liabilities	69,143
	7	Feb-May 2003	Income, Retirement Accts	68,694
	7	Feb-May 2003	Retirement, Pension Plans	68,694
	9	Oct 2003-Jan 2004	Assets, Liabilities	65,901
2004	3	Oct 2004-Jan 2005	Assets, Liabilities	99,978
	4	Feb-May 2005	Income, Retirement Accts	98,098
	6	Oct 2005-Jan 2006	Assets, Liabilities	94,617
	7	Feb-May 2006	Income, Retirement Accts	92,802
	7	Feb-May 2006	Retirement, Pension Plans	92,802
2008	3	May-Aug 2009	Retirement, Pension Plans	95,252
	4	Sep-Dec 2009	Assets, Liabilities	91,219
	5	Jan-Apr 2010	Income, Retirement Accts	90,477
	7	Sep-Dec 2010	Assets, Liabilities	TBD*
	8	Jan-Apr 2011	Income, Retirement Accts	TBD*
	10	Sep-Dec 2011	Assets, Liabilities	TBD*
	11	Jan-Apr 2012	Retirement, Pension Plans	TBD*

*To be determined. Data are expected to be released approximately nine months after the end of the interview period.

The SIPP survey wave files contain an analysis weight variable. For details on the construction of weights, see Westat et al. (2001).

Topical Modules

Annual Income and Retirement Accounts

The Annual Income and Retirement Accounts module collects information on IRA/Keogh accounts and DC plans, among other accounts.

For IRA/Keogh accounts, the survey asks about contributions and withdrawals made during the prior year (as applied to the tax return which is due around the time of the interview). It also asks about account earnings. For example, "Including all IRA accounts in [the respondent's] own name, how much did [his] IRA accounts earn during 2009?" Finally, it asks which types of assets the account holds, distinguishing the following types:

- (1) Certificates of deposit or other savings certificates
- (2) Money market funds
- (3) U.S. Government securities
- (4) Municipal or corporate bonds
- (5) U.S. Savings Bonds
- (6) Stocks or mutual fund shares
- (7) Other assets

The respondent may report multiple asset types but is not asked about the relative size of each type in the account.

Similar questions are asked about the respondent's DC plan(s), where the following asset types are distinguished:

- (1) Money market funds
- (2) U.S. Government securities
- (3) Municipal or corporate bonds
- (4) Stocks or mutual fund shares
- (5) Other assets

Retirement and Pension Plan Coverage

Generally, the Retirement and Pension Plan Coverage module obtains information about the respondent's employer's pension-plan coverage and the receipt of retirement benefits from any former job. Additionally, the module includes questions about pension coverage, vesting, reasons for participation or lack thereof, amount of money in account and how invested; and it includes questions on the pension payments, expected pension payments, loans against retirement and welfare assistance.

The module collects information on plan balances and their asset composition, distinguishing:

- (1) Company stock
- (2) Stock funds
- (3) Corporate bonds or bond funds
- (4) Long term interest bearing securities
- (5) Diversified stock and bond funds
- (6) Government securities
- (7) Money market funds
- (8) Other investments

There is no question on relative allocations of the account balance, although the respondent is asked a follow-up question about the asset type in which the largest share of current contributions are being invested. Respondents are questioned about as many as three DC plans.

Assets and Liabilities

The Assets and Liabilities module also asks about IRA/Keogh accounts and DC plan accounts, but its questions are less detailed than those in the above-discussed modules. Total balances and asset allocations are asked for all IRAs combined, any Keogh accounts, and all DC plans combined. No questions are asked about contributions or withdrawals from retirement accounts.

This module proceeds to ask more detailed questions about the value of various types of after-tax assets, including checking accounts, savings accounts, money market accounts, certificates of deposit, municipal or corporate bonds, U.S. government securities, and stocks or mutual funds (and debt or margin accounts). There are no questions on the institutions which hold the reported assets.

Other Modules

Table 11 lists all topical modules in the 2008 panel.

Table 11. Topical Modules in the 2008 Panel

Wave	Topical modules	Available
1	Reciprocity History Employment History Tax Rebates	YES
2	Work Disability History Education and Training History Marital History Migration History Fertility History Household Relationships Tax Rebates	YES
3	Welfare Reform Retirement and Pension Plan Coverage	
4	Assets and Liabilities Real Estate, Dependent Care, and Vehicles Int Accts, Stocks, Mortg, Val of Bus, Rental, Other Medical Expenses/Utilization of Health Care - Adults and Children Poverty (Work-related Expenses/Child Support Paid) Child Well-Being Economic Stimulus Questions	YES
5	Annual Income and Retirement Accounts Taxes Child Care Work Schedule	YES
6	Adult Well-being Child Support Agreements Support for Non-household Members Functional Limitations and Disability - Adults Functional Limitations and Disability - Children Employer-Provided Health Benefits	YES
7	Assets and Liabilities Real Estate, Dependent Care, and Vehicles Int Accts, Stocks, Mortgage, Val of Bus, Rental, Other Medical Expenses/Utilization of Health Care - Adults and Children Poverty (Worked-related Expenses/Child Support Paid)	Sep 2011
8	Annual Income and Retirement Accounts Taxes Child Care Work Schedule	Jan 2012
9	Informal Care-giving Adult Well-being	May 2012
10	Assets and Liabilities Real Estate, Dependent Care, and Vehicles Int Accts, Stocks, Mortg, Val of Bus, Rental, Other Medical Expenses/Utilization of Health Care - Adults and Children Poverty (Worked-related Expenses/Child Support Paid)	Sep 2012

Wave	Topical modules	Available
11	Child Well-Being Retirement and Pension Plan Coverage	Jan 2013

Source: U.S. Census Bureau, Demographics Survey Division, SIPP branch 2009

Summary Statistics

Tables 12 and 13 tabulate the asset types that were reported for IRAs and DC plans, respectively, in Wave 5 of the 2008 Panel. As explained above respondents may report multiple asset types, but they are not asked for the relative distribution. The figures in the tables therefore do not sum to 100%.

Table 12. Asset Types Reported in IRAs (2008 Panel, Wave 5)

Asset type	Percent of respondents	Weighted percent
Certificates of deposit	14.6%	14.5%
Money market	18.4%	18.6%
U.S. gov't securities	5.8%	6.0%
Municipal/corporate bonds	9.9%	9.9%
U.S Savings Bonds	5.5%	5.7%
Stocks, mutual funds	58.2%	58.2%
Other	7.8%	7.8%

Note: Percentages do not sum to 100% because multiple asset types could be reported.

Table 13. Asset Types Reported in DC Plans (2008 Panel, Wave 5)

Asset type	Percent of respondents	Weighted percent
Money market	19.5%	19.8%
U.S. gov't securities	9.2%	9.2%
Municipal/corporate bonds	13.1%	13.3%
Stocks, mutual funds	59.2%	59.4%
Other	9.1%	9.0%

Note: Percentages do not sum to 100% because multiple asset types could be reported.

Table 14 tabulates demographic characteristics in Wave 3 of the 2008 panel.

Table 14. Demographic Characteristics (2008 Panel, Wave 3)

	Freq.	Percent	Weighted percent
Marital Status			
Married	39,623	41.6%	41.5%
Widowed	4,767	5.0%	4.6%
Divorced	7,830	8.2%	8.1%
Separated	1,452	1.5%	1.5%
Never married	41,580	43.7%	44.3%
Education			
N/A	19,884	20.9%	20.4%
High school or less	11,456	12.0%	11.6%
High school graduate	21,252	22.3%	21.9%
Some college	10,589	11.1%	11.6%
Trade/vocational school	7,946	8.3%	8.2%
Associate's degree	5,734	6.0%	6.2%
Bachelor's degree	11,895	12.5%	13.0%
Master's degree	4,742	5.0%	5.1%
Professional (MD,DDS,JD)	995	1.0%	1.1%
Doctorate	759	0.8%	0.8%
Ethnicity			
Hispanic origin	12,492	13.1%	15.8%
Race			
White	75,360	79.1%	79.8%
Black	11,937	12.5%	12.6%
Asian	3,942	4.1%	3.9%
Other	4,013	4.2%	3.7%
Sex			
Male	45,791	48.1%	49.0%
Female	49,461	51.9%	51.0%
Age			
<18	24,120	25.3%	24.6%
18-25	8,457	8.9%	9.6%
25-35	11,629	12.2%	13.4%
35-45	12,444	13.1%	13.6%
45-55	14,105	14.8%	14.7%
55-65	11,513	12.1%	11.5%
65-75	7,059	7.4%	6.8%
75-85	5,786	6.1%	5.7%
>85	139	0.1%	0.2%

I.4. OTHER DATA SOURCES

This section's facts, statements, and tabulations are current as of September 30, 2011.

Health and Retirement Study

Website: <http://hrsonline.isr.umich.edu>

Collected by the University of Michigan, the HRS is a panel survey of individuals over age 50 and their spouses, regardless of age. It attempts to re-interview respondents every other year. The survey started in 1992. Its sample was augmented in 1998, 2004, and 2010 with individuals from younger cohorts who had become age-eligible. The sample size in the 2010 wave, which is currently available as an early release file, was 15,372 respondents in 10,754 households. In total, about 31,000 respondents have been interviewed.

The HRS asks questions about employment, assets and income, and expectations and risk tolerance, along with a wide range of such other subjects as demographics, health, health care utilization, cognition, marital status, and housing. The survey has been linked to administrative records on Social Security and Medicare.

Respondents who are employed are asked about their participation in a DC plan, their contributions, their account balance, and their asset allocation (percentage in stocks, percentage in employer stock). This information is collected on up to three DC plans for the respondent's current job. Similar information is collected on DC plans sponsored by prior employers. The HRS requests that respondents provide a copy of the Summary Plan Description.

Respondents are also asked about account balance and asset allocation (percentage in stocks or mutual funds) of up to three IRAs or Keogh accounts. In addition, they are asked about contributions to and withdrawals from these accounts since the previous survey interview.

Finally, the HRS asks about after-tax financial instruments. It collects information on the aggregate account balance of and dividend or interest income from (1) stocks or stock mutual funds, (2) bonds or bond mutual funds, (3) checking accounts, savings accounts, or money market accounts, and (4) certificates of deposit, government savings bonds, or Treasury Bills.

Given the limited age range of HRS respondents, the HRS may not be well-suited to derive statistics on asset holdings by the U.S. population as a whole.

Panel Study of Income Dynamics

Website: <http://psidonline.isr.umich.edu>

Collected by the University of Michigan, the PSID is a panel study which began in 1968. It has been collected annually through 1997 and every other year thereafter. The PSID attempts to re-interview its respondents and also interview former household members who split off to form their own household. The 1968 sample of 5,000 households has grown to about 9,000 households in recent waves.

The PSID collects information on a wide range of topics. The core questionnaires ask about demographic information, family structure, health, employment, and income. Since 1999, questions are also asked about aggregate account balances of DC plans, "private annuities and IRAs", and broad categories of after-tax investments. For DC plans, limited information is asked about asset allocation: mostly stocks, mostly in bonds or annuities, or split. Similarly, for private annuities and IRAs, limited information is asked about asset allocation: mostly stocks, mostly interest-earning, or split. Respondents are also asked about regular contributions to their DC plans, possibly as a percentage of pay.

Employee Benefit Research Institute Data

Website: <http://www.ebri.org>

EBRI is an independent, non-profit and non-partisan organization that focuses on employee benefit data and research. It has several sources of data including the Retirement Confidence Survey (RCS), the EBRI IRA Database and a 401(k) database. EBRI, however, does not allow third-party access to its proprietary data.

The proprietary data consist of:

1. RCS:
 - a. Conducted annually in conjunction with American Savings Education Council (ASEC) and Matthew Greenwald & Associates
 - b. Surveys 1,000 people on basic retirement savings information and expectations
2. EBRI IRA Database:
 - a. Collects data from IRA plan administrators
 - b. Contains 14 million accounts with \$732 billion in assets
 - c. Does not currently track changes over time (i.e., withdrawals, contributions and rollovers), but expansion is currently underway
3. EBRI/ICI 401(k) Database
 - a. Conducted with Investment Company Institute (ICI)
 - b. Contains 50,000 plans, 20 million participants and \$1.2 trillion in assets
 - c. Annual updates track balances, asset allocation and loan behavior

The Center for Research in Security Prices Data

Website: <http://www.crsp.com>

CRSP was founded in 1960 as part of the University of Chicago's Booth School of business, specializing in securities prices. CRSP maintains a number of databases including: a database of daily stock prices, dividends and rates of return since 1926; the CRSP/Ziman Real Estate Data Series created with the Ziman Center of the

Anderson School of Business, University of California, Los Angeles; and a database of active and inactive mutual funds created by Mark Carhart of Goldman Sachs Asset Management. CRSP data is available for purchase by subscription only.

The CRSP Survivor-Bias-Free US Mutual Fund Database attempts to provide complete data on both active and inactive mutual funds. The CRSP website states that "the survivor-bias-free nature of the database ensures accurate performance benchmarks and valid analysis." The database contains the following:

- Data on 44,888 funds
 - 17,565 inactive
 - 27,323 active
- Fund Types
 - Equity
 - Taxable and municipal bonds
 - International
 - Money market
- Data beginning in 1962 including:
 - History of each mutual fund's name
 - Investment style
 - Fee structure
 - Holdings
 - Asset allocation
 - Monthly total returns
 - Monthly total net assets (TNA)
 - Monthly/daily net asset values (DNAV/NAV)
 - Distributions
 - Dividends
 - Schedules of rear and front load fees
 - Asset class codes
 - Management company contact information
 - Identifying information
- 2 year contracts at \$14,080 per annum will allow five users access to all data updated quarterly

Financial Research Corporation Data

Website: <http://www.frcnet.com>

FRC compiles data for investment professionals on financial industry trends and financial product performance. The data are at the mutual-fund/variable-annuity level, and they are further broken down by advisor/sponsor and sub-advisor. Data on fees, breakpoints, asset amounts and classes, distribution channel, investment objective, and fund rankings are presented. Manager changes as found in SEC filings are also tracked. Most of these data are measured in bi-annual snapshots.

E-filings of SEC Form N-SAR

Website: <http://www.sec.gov/about/forms/secforms.htm>

On Form N-SAR, registered investment management companies report the operations of their combined share classes on a semi-annual basis. Variables measured include amount of share sales, portfolio turnover rate, front-loads paid by consumers, total inflows subject to loads, and total front- and back-loads paid by consumers. These loads are disaggregated between unaffiliated brokers and captive brokers.

FINRA BrokerCheck Reports

Website: <http://www.finra.org/Investors/ToolsCalculators/BrokerCheck/>

FINRA's BrokerCheck provides investors with information about FINRA registered brokers and brokerage firms. The database contains information from the securities industry's registration and licensing process on 1.3 million current and former brokers and 17,000 brokerage firms. BrokerCheck derives its data from the Central Registration Depository (CRD) which contains required broker licensing forms and regulator-reported information. Brokers and brokerage firms are required to update the CRD within 30 days of an event. Below is a summary of the information contained in BrokerCheck reports.

1. FINRA registered brokerage firm:
 - Report summary with brief overview of the firm and its background
 - Firm profile that describes where and when the firm was founded and lists the people and organizations that own controlling shares or influence daily operations
 - Firm history detailing mergers, acquisitions and name changes
 - Firm operations section that describes the firm's type of business as well as active licenses and registrations
 - Disclosure section listing active and resolved incidents including arbitration awards, disciplinary events and financial matters that are on the firm's record
2. Broker registered within the last 10 years
 - Report summary with overview of broker's credentials
 - Broker qualifications section describing broker's licenses, registrations and completed industry exams
 - Registration and employment section containing two types of information
 - List of the broker's current or previous FINRA registered employers
 - Broker's full ten year employment history in and outside of the securities industry as reported by the broker on his or her Form U4 (submitted by securities firms to states and self-regulatory organizations when registering brokers)
 - Disclosure section that includes customer disputes, disciplinary events and financial matters that have been reported by regulators, involved firms and the broker
3. Broker with registration expired more than 10 years ago:
 - Report summary, broker qualifications and registration and employment sections are the same as a broker registered within the last ten years

- Disclosure section only includes certain civil, criminal, regulatory or judicial complaints. To be included, the broker must have been:
 - Subject of a final regulatory action,
 - Convicted or pled guilty to certain crimes,
 - Subject to investment related civil injunction or found by a civil court to have violated investment related regulations, or
 - Involved in civil litigation or arbitration where the broker was accused of a sales practice violation and judgment went against the broker.

SEC Form ADV

Website: <http://www.adviserinfo.sec.gov/>

Investment advisors file a Form ADV to register with the U.S. Securities and Exchange Commission (SEC) and state securities authorities. This form includes information about the investment advisor's business, ownership, clients, employees, business practices, affiliations, and disciplinary actions against the advisor or any of its employees. Beginning in 2011, a second portion of the form requires investment advisors to prepare non-technical brochures with information on the types of advisory services offered, the advisor's fee schedule, disciplinary information, conflicts of interest, and background information on the management and key advisory personnel.

Some potential areas of interest on the Form ADV are included below.

1. Item 5: Information About Your Advisory Business
 - Approximate number of clients
 - Approximate percentage of clients by client type (i.e., individual, banking institution, charitable organization, etc. IRAs and DC plans are combined under the "individual" category)
 - Approximate percentage of assets by client type
 - Compensation arrangement type, total assets under management and number of accounts
 - Additional information regarding the type of services and activities engaged in by the investment advisor
2. Item 8: Participation or Interest in Client Transactions
 - Attempts to identify potential conflicts of interest by asking questions regarding proprietary and sales interest, and investment or brokerage discretion in client transactions

Some potential questions of interest are included below as they appear on the Form ADV. All of the questions are answered in check box format where Yes or No is the only valid response except when noted.

- A. You are compensated for your investment advisory services by (check all that apply):
 1. A percentage of assets under your management
 2. Hourly Charges
 3. Subscription fees (for a newsletter or periodical)
 4. Fixed fees (other than subscription fees)
 5. Commissions
 6. Performance-based fees

7. Other (specify): _____
- B. Do you or any related person:
1. Buy securities for yourself from advisory clients, or sell securities you own to advisory clients (principal transactions)?
 2. Buy or sell for yourself securities (other than shares of mutual funds) that you also recommend to advisory clients?
 3. Recommend securities (or other investment products) to advisory clients in which you or any related person has some other proprietary (ownership) interest (other than those mentioned in Items 8.A.91) or (2))?
 4. As a broker-dealer or registered representative of a broker-dealer, execute securities trades for brokerage customers in which advisory client securities are sold to or bought from the brokerage customer (agency cross transactions)?
 5. Recommend purchase of securities to advisory clients for which you or any related person serves as underwriter, general or managing partner, or purchaser representative?
 6. Recommend purchase or sale of securities to advisory clients for which you or any related person has any other sales interest (other than the receipt of sales commissions as a broker or registered representative of a broker-dealer)?
- C. Do you or any related person have discretionary authority to determine the:
1. Securities to be bought or sold for a client's account?
 2. Amount of securities to be bought or sold for a client's account?
 3. Broker or dealer to be used for a purchase or sale of securities for a client's account?
 4. Commission rates to be paid to a broker or dealer for a client's securities transactions?
- D. If you answered "yes" to C.(3) above, are any of the brokers or dealers related persons?
- E. Do you or any related person recommend brokers or dealers to clients?
- F. If you answered "yes" to E above, are any of the brokers or dealers related persons?
- G.
1. Do you or any related person receive research or other products or services other than execution from a broker-dealer or third party ("soft dollar benefits") in connection with client securities transactions?
 2. If you answered "yes" to G.(1) above, are all the "soft dollar" benefits you or any related persons receive eligible "research or brokerage services" under section 28(e) of the Securities Exchange Act of 1934?
- H. Do you or any related person, directly or indirectly, compensate any person for client referrals?
- I. Do you or any related person, directly or indirectly, receive compensation from any person for client referrals?

Form ADV data are available in spreadsheet format from the SEC. The Investment Advisor Information Report file contains approximately 11,700 investment advisors as of October 3, 2011.

State Financial Regulators

Website: <http://www.nasaa.org/about-us/contact-us/contact-your-regulator/>

We did not locate any data which would plausibly assist this Task Order.

I.5. BIBLIOGRAPHY

- Bricker, Jesse, Brian Bucks, Arthur Kennickell, Traci Mach, and Kevin Moore (2011). "Surveying the Aftermath of the Storm: Changes in Family Finances from 2007 to 2009." Board of Governors of the Federal Reserve System. Retrieved from <http://www.federalreserve.gov/pubs/feds/2011/201117/201117pap.pdf>.
- Kennickell, Arthur B. (1998). "Multiple Imputation in the Survey of Consumer Finances." Board of Governors of the Federal Reserve System. Retrieved from <http://www.federalreserve.gov/pubs/oss/oss2/papers/impute98.pdf>.
- Kennickell, Arthur B. (1999). "Revisions to the SCF Weighting Methodology: Accounting for Race/Ethnicity and Homeownership." Board of Governors of the Federal Reserve System. Retrieved from <http://www.federalreserve.gov/pubs/oss/oss2/papers/weight.revision.pdf>.
- Kennickell, Arthur B. (2000). "Revisions to the Variance Estimation Procedure for the SCF." Board of Governors of the Federal Reserve System. Retrieved from <http://www.federalreserve.gov/pubs/oss/oss2/papers/variance.pdf>.
- Kennickell, Arthur B. and R. Louise Woodburn (1997). "Consistent Weight Design for the 1989, 1992 and 1995 SCFs, and the Distribution of Wealth." Board of Governors of the Federal Reserve System. Retrieved from <http://www.federalreserve.gov/pubs/oss/oss2/papers/wgt95.pdf>.
- SCF Codebook (2001). "Codebook for 2001 Survey of Consumer Finances." Board of Governors of the Federal Reserve System, Washington, DC. Retrieved from <http://www.federalreserve.gov/econresdata/scf/files/codebk2001.txt>.
- SCF Codebook (2004). "Codebook for 2004 Survey of Consumer Finances." Board of Governors of the Federal Reserve System, Washington, DC. Retrieved from <http://www.federalreserve.gov/pubs/oss/oss2/2004/codebk2004.txt>.
- SCF Codebook (2007). "Codebook for 2007 Survey of Consumer Finances." Board of Governors of the Federal Reserve System, Washington, DC. Retrieved from <http://www.federalreserve.gov/econresdata/scf/files/codebk2007.txt>.
- SCF Codebook (2010). "Codebook for 2010 Survey of Consumer Finances." Board of Governors of the Federal Reserve System, Washington, DC. Retrieved from <http://www.federalreserve.gov/econresdata/scf/files/codebk2010.txt>.
- SCF Codebook (2013). "Codebook for 2013 Survey of Consumer Finances." Board of Governors of the Federal Reserve System, Washington, DC. Retrieved from <http://www.federalreserve.gov/econresdata/scf/files/codebk2013.txt>.
- U.S. Census Bureau, Demographics Survey Division, Survey of Income and Program Participation branch (2009). "Overview of the Survey of Income and Program Participation (SIPP)." Retrieved from <http://www.census.gov/sipp/index.html>.
- Westat and Mathematica Policy Research Inc. (2001). "Survey of Income and Program Participation Users' Guide. Third Edition." Washington, DC. Retrieved from <http://www.census.gov/sipp/usrguide/sipp2001.pdf>.

1.1. INTRODUCTION

The first part of the report is devoted to a general introduction of the subject. It starts with a brief history of the field, followed by a discussion of the current state of research. The main objectives of the study are then outlined, and the structure of the report is explained. The second part of the report is devoted to the statistical analysis of the data. This part is divided into two main sections: the first section deals with the description of the data, and the second section deals with the inference. The third part of the report is devoted to the conclusions and recommendations. This part starts with a summary of the main findings, followed by a discussion of the implications of these findings. Finally, some recommendations for further research are given.

PART II. STATISTICAL ANALYSIS

The statistical analysis is divided into two main parts: the first part deals with the description of the data, and the second part deals with the inference. In the first part, the distribution of the data is described, and the main characteristics of the data are summarized. In the second part, the inference is carried out, and the results are compared with the theoretical expectations. The results of the statistical analysis are presented in the following sections.

The results of the statistical analysis show that the data are normally distributed, and that the mean and standard deviation are as expected. The inference shows that the results are significant, and that the null hypothesis is rejected. The conclusions and recommendations are given in the following sections.

II.1. INTRODUCTION

Part II of this report compiles the analyses and statistical tabulations completed for the "SCF Statistical Studies and Literature Analysis" task order. This includes tabulations and analyses relating to the differences among owners of individual retirement accounts (IRAs), 401(k) and other defined contribution (DC) plans, and after-tax investment accounts.

The IRA category includes:

- Roth IRAs,
- IRAs rolled over from pension accounts,
- Regular or other IRAs, and
- Keogh Accounts.

The DC category includes:

- 401(k),
- 403(b),
- 457,
- Thrift/Savings,
- Supplemental Retirement Annuities (SRA),
- Simplified Employee Pensions (SEP),
- Simplified Incentive Match Plans for Employers (SIMPLE),
- Money Purchase Plans,
- Other current job plans from which loans or withdrawals can be made, and
- Accounts from past jobs from which the family expects to receive the account balance in the future or is currently drawing from the account. (Note: The DC category excludes cash balance plans.³)

The taxable asset category includes:⁴

- Mutual funds,
- Bonds (other than savings bonds),
- Publicly traded stock,
- Cash or call-money accounts held at brokerages, and
- Trusts that can be cashed in.

As requested by staff of the U.S. Department of Labor's Employee Benefits Security Administration (EBSA), all data in this document are sourced from the 2013 Survey of Consumer Finances (SCF) and weighted by household weights unless otherwise stated. Note that households with IRAs, DC plans, and taxable accounts are not mutually exclusive classifications of households.

³ The account type of plans from prior jobs is unknown. Some cash balance plans from prior jobs may be included in the analysis. Also, cash balance plans on the current job from which the respondent can borrow or withdraw may be included.

⁴ The taxable asset category does not include savings accounts, certificates of deposit (CDs), cashable life insurance, cashable annuities or savings bonds. However, the document does analyze these after-tax investment types individually when noted.

II.2. TABULATIONS AND ANALYSES

Table 15. Categories of Households by Ownership and Value of Account

	Households		Assets	
	Number of households (millions)	Percent	Total value of accounts (\$ billions)	Percent
<i><u>Not mutually exclusive:</u></i>				
IRA	34.4	28.1%	19,557.5	80.2%
DC	43.2	35.2%	12,897.7	52.9%
Taxable	24.3	19.9%	19,910.8	81.7%
<i><u>Mutually exclusive:</u></i>				
None	56.5	46.1%	0.0	0.0%
IRA only	10.9	8.9%	1,280.3	5.3%
DC only	22.3	18.2%	1,358.6	5.6%
Taxable only	4.8	3.9%	1,507.8	6.2%
IRA & DC	8.5	7.0%	1,831.9	7.5%
IRA & Taxable	7.2	5.9%	8,695.8	35.7%
DC & Taxable	4.5	3.7%	1,957.7	8.0%
IRA, DC & Taxable	7.8	6.4%	7,749.5	31.8%
Total	122.5	100.0%	24,381.6	100.0%

Source: SCF 2013.

Note: Weighted by household weights. Not-mutually-exclusive asset figures include assets from any of the three types. For example, 28.1% of households own an IRA. Their total assets holdings, including DC plan and taxable account balances, sum to \$19.6 trillion, 80.2% of all assets.

Table 16 replicates Table 15, but uses an expanded definition of taxable assets that we refer to as "retail" assets. In addition to the categories of assets that were included in the taxable category (publicly traded stocks, mutual funds, bonds, trusts, and brokerage call (cash) accounts) the retail asset category also includes savings accounts, CDs, cashable life insurance, cashable annuities, and savings bonds.

**Table 16. Categories of Households by Ownership and Value of Accounts
(including all retail investments and savings)**

	Households		Assets	
	Number of households (millions)	Percent	Total value of accounts (\$ billions)	Percent
<i>Not mutually exclusive:</i>				
IRA	34.4	28.1%	22,428.5	78.6%
DC	43.2	35.2%	14,734.7	51.6%
Retail*	83.6	68.2%	27,846.1	97.5%
<i>Mutually exclusive:</i>				
None	29.2	23.8%	0.0	0.0%
IRA Only	2.4	1.9%	207.0	0.7%
DC Only	6.2	5.0%	270.2	0.9%
Retail Only	32.1	26.2%	2,274.3	8.0%
IRA & DC Only	1.2	1.0%	227.2	0.8%
IRA & Retail Only	15.7	12.8%	11,334.5	39.7%
DC & Retail Only	20.6	16.8%	3,577.6	12.5%
IRA, DC & Retail	15.2	12.4%	10,659.7	37.3%
Total	122.5	100.0%	28,550.5	100.0%

Source: SCF 2013.

*Retail financial assets include publicly traded stocks, mutual funds, bonds, trusts, brokerage call (cash) accounts, savings accounts, CDs, cashable life insurance, cashable annuities, and savings bonds.

Note: Weighted by household weights. Not-mutually-exclusive asset figures include assets from any of the three types. For example, 28.1% of households own an IRA. Their total assets holdings, including DC plan and retail account balances, sum to \$22.4 trillion, 78.6% of all assets.

Table 17. Taxable Asset Holders and Holdings by Asset Type

Asset type	Households holding assets (millions)	Percent of all households	Total balance (\$ billions)	Percent of taxable assets
Stock	16.8	13.8%	4,792.3	38.7%
Mutual fund	10.0	8.2%	4,548.7	36.7%
Bond	1.7	1.4%	959.3	7.7%
Trust	1.8	1.5%	1,687.6	13.6%
Brokerage call	2.4	1.9%	409.8	3.3%
Total			12,397.6	100.0%

Source: SCF 2013.

Note: Weighted by household weights. Holders of taxable asset types are not mutually exclusive. Trust holdings include cash-out eligible accounts only.

Table 18. Income Distribution for All Households and for Those with IRAs, DC Plans, or Taxable Accounts

Annual household income	All households		Households with IRAs		Households with DC plans		Households with taxable accounts	
	Number (millions)	Percent	Number (millions)	Percent	Number (millions)	Percent	Number (millions)	Percent
Under \$10,000	3.6	2.9%	0.1	0.3%	0.1	0.2%	0.1	0.6%
\$10,000 - \$20,000	14.9	12.1%	0.5	1.4%	0.6	1.3%	0.6	2.4%
\$20,000 - \$50,000	41.9	34.2%	6.7	19.4%	8.6	20.0%	4.4	18.0%
\$50,000 - \$100,000	35.3	28.8%	11.4	33.2%	16.5	38.3%	6.9	28.5%
\$100,000 - \$200,000	19.2	15.7%	10.1	29.3%	12.3	28.6%	7.2	29.5%
\$200,000 - \$500,000	5.9	4.8%	4.2	12.3%	3.7	8.7%	3.5	14.3%
\$500,000 and above	1.9	1.6%	1.4	4.1%	1.3	3.0%	1.6	6.6%
Total	122.5	100.0%	34.4	100.0%	43.2	100.0%	24.3	100.0%

Source: 2013 SCF.

Note: Weighted by household weights. Income includes withdrawals from DC plans and IRAs.

Figure 1. Income Distribution for All Households and for Those with IRAs, DC Plans, or Taxable Accounts

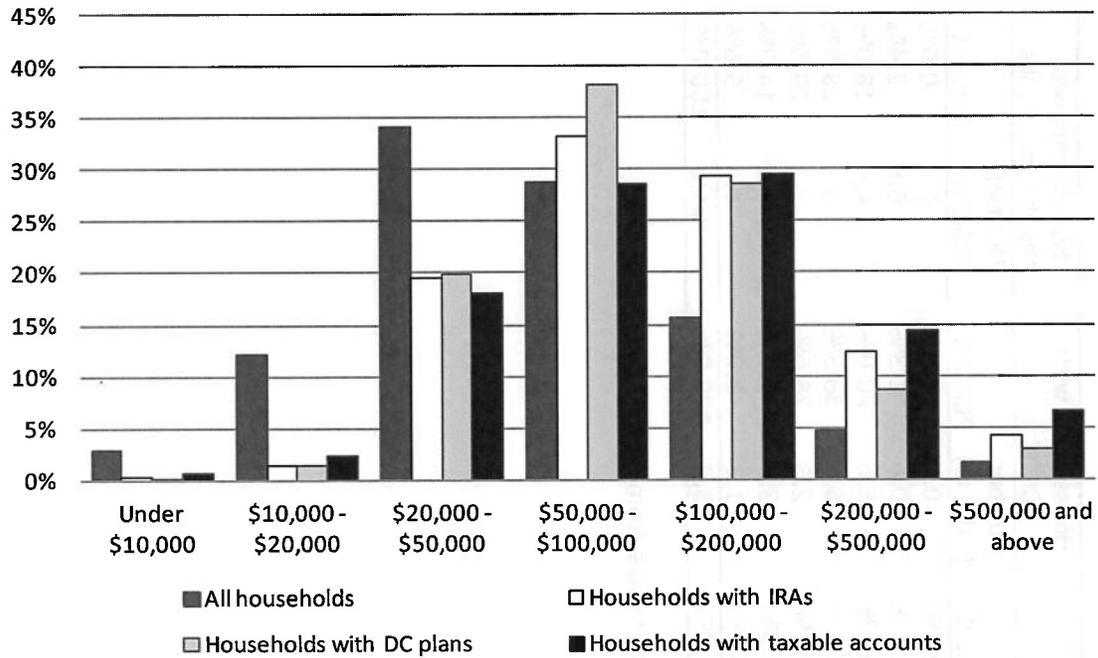


Table 19. Income Distribution of Asset Holdings for Households with IRAs, DC Plans, or Taxable Accounts

Annual household income	Households with IRAs		Households with DC plans		Households with taxable accounts	
	Value of accounts (\$ billions)	Percent	Value of accounts (\$ billions)	Percent	Value of accounts (\$ billions)	Percent
Under \$10,000	4.1	0.1%	0.3	0.0%	19.4	0.2%
\$10,000 - \$20,000	26.3	0.4%	9.1	0.2%	71.4	0.6%
\$20,000 - \$50,000	458.2	6.9%	191.3	3.6%	464.8	3.7%
\$50,000 - \$100,000	1,390.4	20.8%	1,083.5	20.4%	1,132.1	9.1%
\$100,000 - \$200,000	1,857.8	27.8%	1,924.3	36.3%	1,888.5	15.2%
\$200,000 - \$500,000	1,712.3	25.6%	1,124.0	21.2%	3,178.2	25.6%
\$500,000 and above	1,226.9	18.4%	975.6	18.4%	5,643.4	45.5%
Total	6,676.0	100.0%	5,308.0	100.0%	12,397.6	100.0%

Source: 2013 SCF.

Note: Weighted by household weights. Income includes withdrawals from DC plans and IRAs.

Figure 2. Household Income Distribution of Asset Holdings for Households with IRAs, DC Plans, or Taxable Accounts

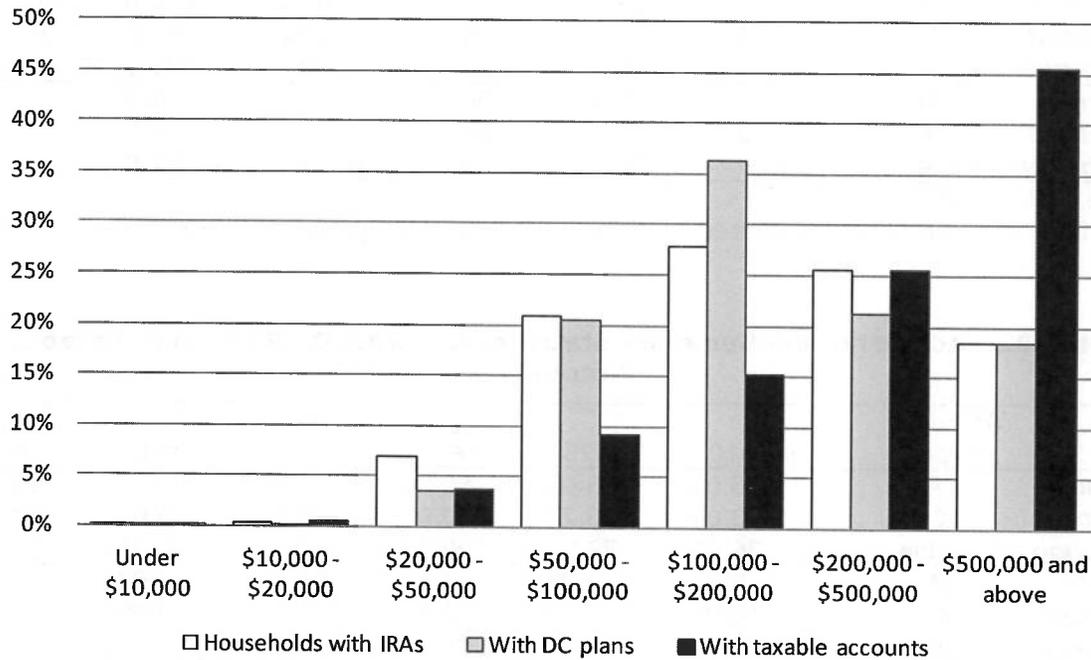


Table 20. Income Summary Statistics for Households with IRAs, DC Plans, or Taxable Accounts (\$ thousands)

Account type		p10	p25	Median	p75	p90	Mean
<i>Household-weighted distribution</i>	IRA	35.5	54.7	93.3	153.2	269.9	153.0
	DC	34.5	52.8	85.2	133.9	213.1	129.8
	Taxable	31.5	55.8	100.4	181.6	329.7	185.0
	All HHs	15.2	27.4	50.7	91.3	154.9	85.7
<i>Asset-weighted distribution</i>	IRA	54.8	90.3	175.5	355.1	760.9	
	DC	67.0	101.5	157.3	322.6	759.9	
	Taxable	81.2	175.5	396.7	1,148.5	2,800.1	

Source: 2013 SCF.

Note: The n-th percentile is denoted "pn". Income includes withdrawals from DC plans and IRAs. The row labeled "All HHs" reflects all households, including households without any accounts.

Table 21. Account-Balance Summary Statistics for IRAs, DC Plans, and Taxable Accounts (\$ thousands)

Account type		p10	p25	Median	p75	p90	Mean
<i>Household-weighted distribution</i>	IRA	4.0	13.0	50.0	164.0	450.0	193.9
	DC	1.3	8.0	35.0	125.0	300.0	123.0
	Taxable	2.0	9.0	48.0	260.0	975.0	509.2
<i>Asset-weighted distribution</i>	IRA	100.0	261.0	700.0	1,830.0	4,504.0	
	DC	72.0	170.0	366.0	833.0	2,419.0	
	Taxable	350.0	1,080.0	3,525.0	15,450.0	50,799.0	

Source: 2013 SCF.

Note: The n-th percentile is denoted "pn". Accounts are aggregated by households.

Table 22. Stock-Fraction Summary Statistics for IRAs, DC Plans, and Taxable Accounts

Account type		p10	p25	Median	p75	p90	Mean
<i>Household-weighted distribution</i>	IRA	0.0%	15.0%	50.0%	81.6%	100.0%	49.4%
	DC	0.0%	18.2%	50.0%	80.0%	100.0%	49.3%
	Taxable	36.9%	79.9%	100.0%	100.0%	100.0%	84.1%
<i>Asset-weighted distribution</i>	IRA	5.0%	30.0%	54.5%	80.0%	100.0%	
	DC	10.0%	30.0%	52.7%	80.0%	100.0%	
	Taxable	23.1%	47.4%	72.1%	95.7%	100.0%	

Source: 2013 SCF.

Note: The n-th percentile is denoted "pn." Fractions are calculated after aggregating account balances by households.

Table 23. Distribution of Demographic Characteristics of the Heads of All Households and of Those with IRAs, DC Plans, or Taxable Accounts

	All households	With IRAs	With DC plans	With taxable accounts
<u>Marital status</u>				
Married	48.1%	64.7%	62.9%	64.7%
Separated	3.5%	1.0%	1.8%	0.9%
Divorced	16.6%	13.6%	14.7%	10.9%
Widowed	9.6%	8.2%	2.6%	8.4%
Never married	22.1%	12.5%	18.0%	15.0%
<u>Education</u>				
Less than high school	11.0%	1.5%	3.6%	1.9%
High school or equivalency	31.3%	19.7%	24.1%	15.5%
Some college	18.9%	14.1%	17.7%	13.8%
Associate's degree	6.6%	7.1%	8.4%	4.7%
Bachelor's degree	19.7%	31.4%	27.4%	34.4%
Master's degree	9.3%	18.4%	13.6%	20.4%
Professional degree/doctorate	3.2%	7.7%	5.3%	9.3%
<u>Ethnicity</u>				
Hispanic/Latino origin	13.3%	4.8%	8.5%	3.8%
Non-Hispanic/Latino	86.7%	95.2%	91.5%	96.2%
<u>Race</u>				
White	70.1%	86.8%	76.8%	88.0%
Black	14.6%	5.6%	11.8%	3.9%
Hispanic	10.6%	2.7%	6.4%	2.1%
Other	4.7%	4.9%	5.0%	6.0%
<u>Age</u>				
18-25	6.2%	0.7%	3.6%	2.4%
26-35	16.2%	10.7%	19.4%	10.7%
36-45	17.5%	15.9%	23.2%	15.8%
46-55	19.7%	20.2%	26.2%	20.3%
56-65	18.6%	25.9%	21.1%	21.9%
66-75	12.2%	17.2%	5.0%	16.6%
76-85	7.5%	7.6%	1.2%	9.8%
>85	2.3%	1.7%	0.3%	2.5%

Source: 2013 SCF.

Note: Weighted by household weights.

Table 24. Asset-Weighted Distribution of Demographic Characteristics of the Heads of Households with IRAs, DC Plans, or Taxable Accounts

	Households with IRAs	With DC plans	With taxable accounts
<i><u>Marital status</u></i>			
Married	81.0%	80.3%	79.6%
Separated	0.3%	0.8%	1.2%
Divorced	8.3%	10.1%	6.8%
Widowed	6.5%	1.6%	6.4%
Never married	3.9%	7.2%	6.0%
<i><u>Education</u></i>			
Less than high school	0.2%	1.1%	1.1%
High school or equivalency	8.8%	12.5%	5.2%
Some college	9.7%	12.3%	7.6%
Associate's degree	2.7%	5.1%	3.2%
Bachelor's degree	29.8%	31.6%	31.2%
Master's degree	24.9%	20.0%	30.8%
Professional degree/doctorate	23.9%	17.4%	20.9%
<i><u>Ethnicity</u></i>			
Hispanic/Latino origin	2.3%	4.0%	2.3%
Non-Hispanic/Latino	97.7%	96.0%	97.7%
<i><u>Race</u></i>			
White	95.4%	88.2%	95.4%
Black	1.3%	4.6%	0.7%
Hispanic	0.6%	2.2%	0.7%
Other	2.7%	5.0%	3.2%
<i><u>Age</u></i>			
18-25	0.0%	0.3%	0.4%
26-35	1.4%	5.0%	2.7%
36-45	6.3%	17.2%	7.3%
46-55	13.5%	32.1%	20.1%
56-65	30.4%	34.1%	27.9%
66-75	36.7%	9.8%	26.5%
76-85	10.0%	1.3%	10.4%
>85	1.8%	0.2%	4.9%

Source: 2013 SCF.

Note: Weighted by household weights and account balances.

Table 25. Margin Loan Statistics

	p10	p25	Median	p75	p90	Mean
Loan balance (\$ thousands)	4.0	10.0	25.0	250.0	500.0	192.8
Ratio of loan balance to taxable assets	0.3%	6.9%	14.7%	16.7%	40.0%	16.7%

Source: 2013 SCF.

Note: 0.75% of households with taxable assets (0.15% of all households) reported having a margin loan. The n-th percentile is denoted by "pn".
Weighted by household weights.

Table 26: IRAs, IRA assets, and Median Household Income of IRA Holders, by Type of Institution Holding IRAs⁵

	Fraction of IRAs held in institution	Median income of households with account at institution (\$ thousands)	Fraction of IRA assets held in institution	Asset-weighted median income of households with acct at institution (\$ thousands)
Commercial bank	46.0%	90.3	45.8%	164.4
Savings and loan	2.9%	71.0	1.2%	60.9
Credit union	6.3%	71.0	2.2%	72.0
Finance or loan company	1.4%	111.6	1.1%	120.7
Brokerage	40.1%	106.5	45.6%	182.6
Insurance company	1.6%	97.4	0.8%	152.2
Investment/mgmt company	0.9%	93.3	2.6%	335.8
Other	1.0%	74.1	0.8%	101.5
All	100.0%		100.0%	

Source: 2013 SCF.

⁵ Please see Appendix 2 for detailed information on the method used to calculate the numbers in Table 25.

Table 27. Distribution of Asset Holdings By Household Income

	Under \$10,000		\$10,000-\$20,000		\$20,000-\$50,000		\$50,000-\$100,000	
	\$ billions	Percent	\$ billions	Percent	\$ billions	Percent	\$ billions	Percent
DC accounts	0.3	0.9	9.1	4.8	191.3	11.6	1,083.5	23.9
IRAs	4.1	11.6	26.3	13.9	458.2	27.8	1,390.4	30.6
Stocks	16.0	45.5	21.2	11.2	201.4	12.2	533.0	11.7
Mutual funds	2.9	8.3	39.0	20.6	147.2	8.9	392.3	8.6
Savings accounts	8.6	24.4	23.2	12.3	210.4	12.8	374.3	8.2
Bonds	0.2	0.5	1.2	0.7	34.4	2.1	54.9	1.2
CDs	0.1	0.4	35.5	18.8	120.4	7.3	138.3	3.0
Trusts	0.0	0.0	9.8	5.2	66.7	4.0	96.2	2.1
Cashable life	2.0	5.6	11.0	5.8	88.5	5.4	177.0	3.9
Cashable annuities	0.7	1.9	9.3	4.9	100.2	6.1	225.2	5.0
Broker cash	0.3	0.8	0.0	0.0	15.0	0.9	55.7	1.2
Savings Bonds	0.0	0.1	3.5	1.8	15.9	1.0	20.2	0.4
Total	35.1	100.0	189.2	100.0	1,649.8	100.0	4,541.0	100.0

	\$100,000-\$200,000		\$200,000-\$500,000		\$500,000+		Total	
	\$ billions	Percent	\$ billions	Percent	\$ billions	Percent	\$ billions	Percent
DC accounts	1,924.3	29.6	1,124.0	16.5	975.6	11.0	5,308.0	18.6
IRAs	1,857.8	28.6	1,712.3	25.2	1,226.9	13.9	6,676.0	23.4
Stocks	842.9	13.0	1,065.6	15.7	2,112.3	23.9	4,792.3	16.8
Mutual funds	694.6	10.7	1,220.2	17.9	2,052.3	23.2	4,548.7	15.9
Savings accounts	432.1	6.6	426.2	6.3	502.1	5.7	1,977.0	6.9
Bonds	68.6	1.1	369.4	5.4	430.5	4.9	959.3	3.4
CDs	91.0	1.4	85.0	1.2	141.5	1.6	612.0	2.1
Trusts	252.5	3.9	395.2	5.8	867.0	9.8	1,687.6	5.9
Cashable life	137.7	2.1	151.7	2.2	260.7	3.0	828.7	2.9
Cashable annuities	150.7	2.3	108.0	1.6	78.6	0.9	672.7	2.4
Broker cash	29.9	0.5	127.7	1.9	181.2	2.1	409.8	1.4
Savings Bonds	19.6	0.3	14.5	0.2	4.9	0.1	78.6	0.3
Total	6,501.7	100.0	6,799.9	100.0	8,833.7	100.0	28,550.5	100.0

Source: 2013 SCF.

Note: Weighted by household weights. Income includes withdrawals from DC plans and IRAs. Totals differ from those in Table 18 because Table 18 excludes savings accounts, CDs, cashable life insurance, cashable annuities, and savings bonds.

Figure 3. Household Asset Holdings, Income < \$10K

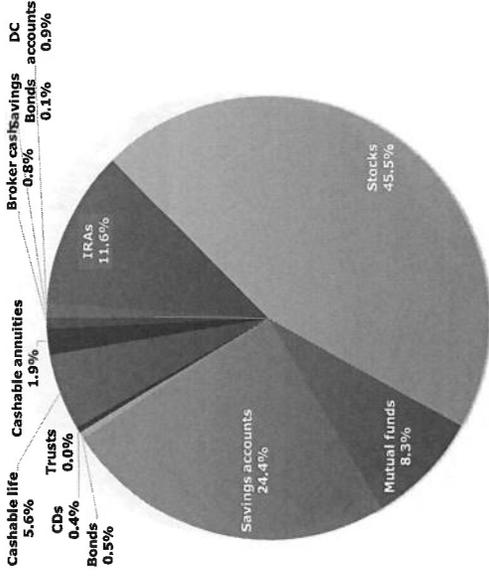


Figure 5. Household Asset Holdings, Income \$20K-50K

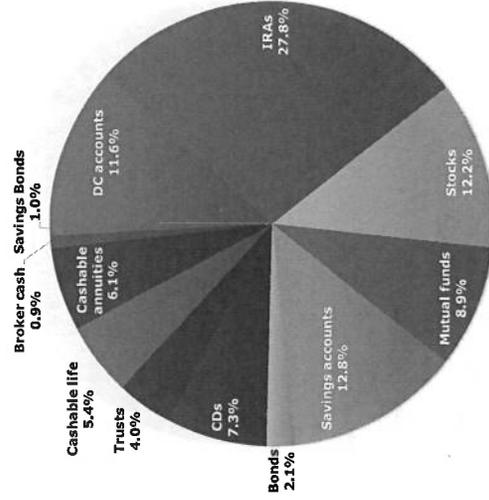


Figure 4. Household Asset Holdings, Income \$10K-20K

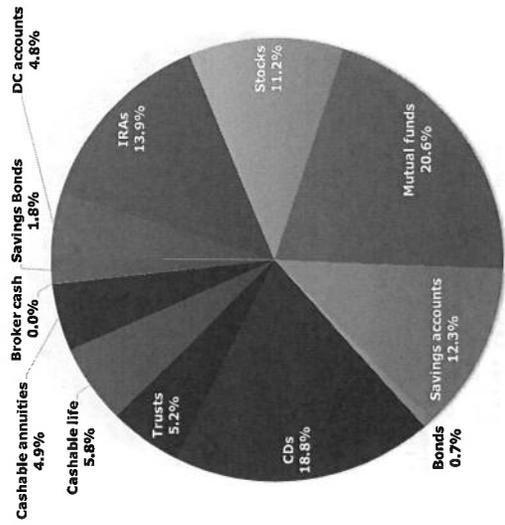


Figure 6. Household Asset Holdings, Income \$50K-100K

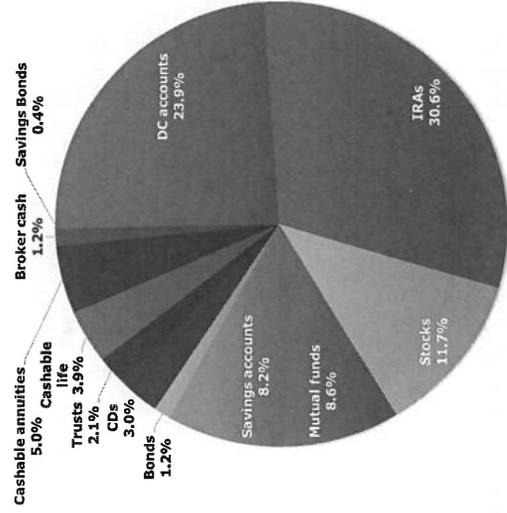


Figure 7. Household Asset Holdings, Income \$100K-200K

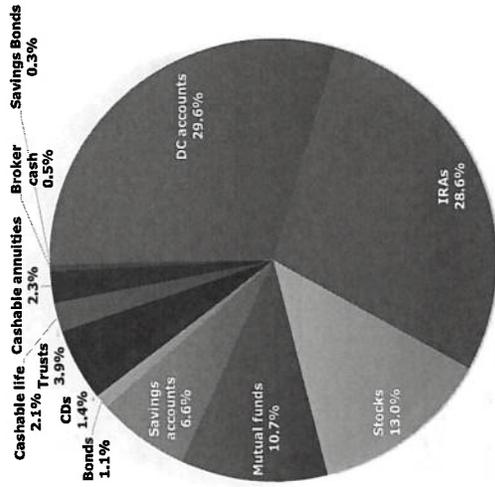


Figure 9. Household Asset Holdings, Income \$500K+

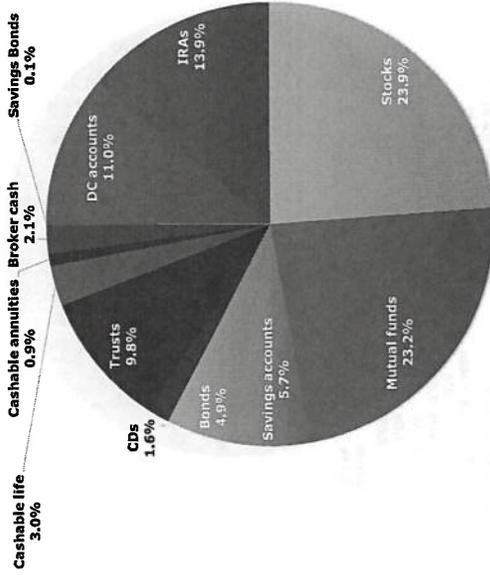


Figure 8. Household Asset Holdings, Income \$200K-500K

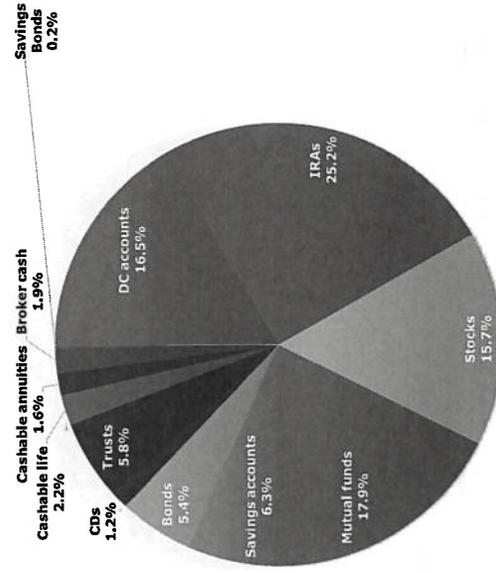


Figure 10. Household Asset Holdings, Total Population

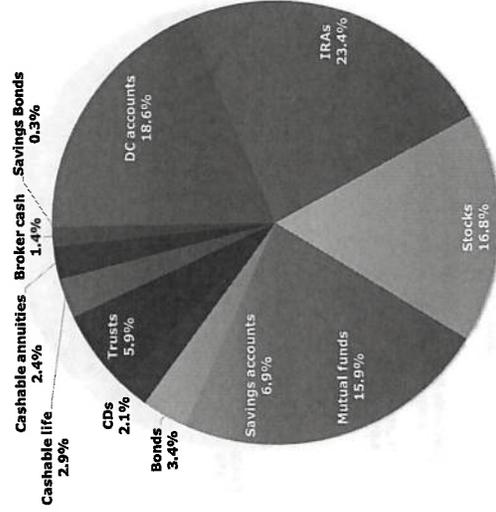


Table 28. Distribution of Households, IRA Accounts, and IRA Assets by Total Household IRA Assets

Household IRA Assets	Households			IRA Accounts			IRA Assets		
	Number of households (millions)	Percent	Cumulative Percent	Number of accounts (millions)	Percent	Cumulative Percent	Value of accounts (\$ billions)	Percent	Cumulative Percent
None	88.1	71.9%	71.9%	0.0	0.0%	0.0%	0.0	0.0%	0.0%
\$1-\$9,999	6.4	5.3%	77.2%	7.6	11.5%	11.5%	25.4	0.4%	0.4%
\$10,000-\$24,999	5.9	4.8%	82.0%	8.6	13.0%	24.5%	93.4	1.4%	1.8%
\$25,000-\$49,999	4.2	3.5%	85.5%	7.3	11.1%	35.6%	149.0	2.2%	4.0%
\$50,000-\$99,999	5.6	4.6%	90.0%	10.7	16.2%	51.8%	388.8	5.8%	9.8%
\$100,000-\$249,999	5.9	4.8%	94.8%	13.5	20.4%	72.1%	912.1	13.7%	23.5%
\$250,000-\$499,999	3.2	2.6%	97.4%	8.2	12.4%	84.5%	1,100.5	16.5%	40.0%
\$500,000-\$999,999	1.9	1.5%	99.0%	5.7	8.6%	93.2%	1,273.9	19.1%	59.1%
\$1,000,000+	1.3	1.0%	100.0%	4.5	6.8%	100.0%	2,733.0	40.9%	100.0%
Total	122.5	100.0%		66.1	100.0%		6,676.0	100.0%	

Source: 2013 SCF.

Note: Weighted by household weights.

Table 29. Distribution of Number of IRA Accounts per Household Member by Total Household IRA Assets

Household IRA Assets	Respondent		Spouse		Other member	
	Total number of accounts (millions)	Percent	Total number of accounts (millions)	Percent	Total number of accounts (millions)	Percent
None	0.0	0.0%	0.0	0.0%	0.0	0.0%
\$1-\$9,999	5.2	12.0%	2.4	10.9%	0.0	1.2%
\$10,000-\$24,999	5.8	13.4%	2.7	12.0%	0.1	17.5%
\$25,000-\$49,999	4.9	11.3%	2.2	9.9%	0.2	37.6%
\$50,000-\$99,999	6.8	15.8%	3.8	17.0%	0.0	7.5%
\$100,000-\$249,999	9.0	20.9%	4.4	19.5%	0.1	11.6%
\$250,000-\$499,999	4.9	11.4%	3.2	14.5%	0.1	10.2%
\$500,000-\$999,999	3.6	8.3%	2.1	9.2%	0.1	11.1%
\$1,000,000+	2.9	6.8%	1.5	6.9%	0.0	3.2%
Total	43.1	100.0%	22.3	100.0%	0.6	100.0%

Source: 2013 SCF.

Note: Weighted by household weights.

Table 30. Household Account Types by Source of Financial Advice

Source	Call around	Magazines/ Newspapers/ Books	Material in the mail	Television/ Radio	Internet/ Online services	Adverts	Friends/ Relatives	Lawyers	Accountant	Bankers	Brokers	Financial planner	Self
Sample Size (n/5)	776.2	817.2	361.8	480.4	2,253.8	356.6	2,313.0	314.8	817.2	2,016.4	765.8	1,806.8	487.6
HHS with IRAs	27.6%	47.3%	31.7%	28.7%	35.4%	25.3%	25.7%	26.3%	38.7%	26.1%	56.4%	46.8%	33.2%
HHS with DC plans	34.0%	36.8%	32.8%	32.1%	46.8%	31.2%	39.6%	26.0%	38.0%	35.2%	42.4%	46.4%	39.9%
HHS with taxable acct	18.9%	37.0%	18.6%	23.4%	26.7%	19.9%	21.2%	22.0%	31.4%	17.3%	44.9%	34.6%	17.1%
HHS without IRAs	72.4%	52.7%	68.3%	71.3%	64.6%	74.7%	74.3%	73.7%	61.3%	73.9%	43.6%	53.2%	66.8%
HHS without DC plans	66.0%	63.2%	67.2%	67.9%	53.2%	68.8%	60.4%	74.0%	62.0%	64.8%	57.6%	53.6%	60.1%

Source: 2013 SCF.

Note: Weighted by household weights. Percentages do not add to 100% because multiple answers were allowed. Sample size contains decimal values because the SCF stores each household response as five replicate values and imputes missing values using five different assumptions; therefore, a household can be imputed into multiple categories of a single variable. For example, a household may hold an IRA in two replications and not hold an IRA in three replications.

Table 31 shows sources of financial advice by various account types that households own or do not own. Its top row shows unweighted sample sizes. The SCF stores each household response as five records ("replicates") and imputes missing values using five different assumptions; therefore, a household can be imputed into multiple categories of a single variable. For example, a household may hold an IRA in two replicates and not hold an IRA in three replicates. The sample sizes reported in Table 31 and subsequent tables are equal to one-fifth of the number of records that match the category, rounded to the nearest integer.

Table 31. Sources of Financial Advice by Household Account Types

Source	All households	Households with IRAs			Households with DC plans		Households with taxable accounts		Households without IRAs		Households without DC plans	
		6,015	2,170	12.9%	2,309	12.7%	1,815	12.5%	3,845	13.3%	3,706	13.4%
Call around	13.2%	12.9%	12.7%	12.5%	13.3%	13.4%	10.7%	6.6%	8.3%	28.9%	6.6%	37.4%
Magazines/Newspapers/Books	11.0%	18.5%	11.5%	20.4%	8.0%	6.1%	7.9%	31.7%	6.4%	41.4%	4.4%	9.7%
Material in the mail	6.4%	7.2%	5.9%	6.0%	7.9%	31.7%	6.4%	41.4%	4.4%	9.7%	33.0%	7.7%
Television/Radio	8.0%	8.1%	7.2%	9.4%	7.9%	31.7%	6.4%	41.4%	4.4%	9.7%	21.1%	6.3%
Internet/Online services	35.3%	44.4%	46.9%	47.5%	31.7%	28.9%	6.6%	37.4%	4.4%	9.7%	53.2%	
Advertisements	6.2%	5.6%	5.5%	6.2%	6.4%	41.4%	4.4%	9.7%	33.0%	7.7%		
Friends/Relatives	40.1%	36.7%	45.0%	42.7%	41.4%	37.4%	4.4%	9.7%	33.0%	7.7%		
Lawyers	3.8%	3.6%	2.8%	4.2%	3.9%	4.4%	9.7%	33.0%	7.7%			
Accountants	10.2%	14.0%	11.0%	16.1%	8.7%	9.7%	33.0%	7.7%	21.1%	6.3%		
Bankers	33.0%	30.7%	32.9%	28.7%	33.9%	33.0%	7.7%	21.1%	6.3%			
Brokers	8.6%	17.3%	10.4%	19.5%	5.2%	21.1%	6.3%					
Financial planners	25.5%	42.4%	33.6%	44.3%	18.8%	6.3%						
Self	6.8%	8.0%	7.7%	5.8%	6.3%							
Financial planner, banker, broker, accountant, or lawyer	56.5%	72.3%	62.5%	73.5%	50.3%							

Source: 2013 SCF.

Note: Weighted by household weights. Percentages do not add to 100% because multiple answers were allowed.

Table 32. Sources of Financial Advice for Households with an IRA, by Household IRA Assets

Sample size	Households with IRAs	Household IRA Assets									
		\$1-\$9,999	\$10,000-\$24,999	\$25,000-\$49,999	\$50,000-\$99,999	\$100,000-\$249,999	\$250,000-\$499,999	\$500,000-\$999,999	\$1,000,000+		
	2,170	296	275	217	308	382	251	205	237		
Call around	12.9%	16.2%	15.5%	12.2%	13.5%	11.9%	9.8%	6.7%	5.8%		
Magazines/Newspapers/Books	18.5%	11.6%	12.8%	18.7%	15.0%	22.0%	29.3%	26.0%	39.3%		
Material in the mail	7.2%	7.8%	6.5%	10.7%	5.1%	7.5%	7.8%	5.1%	4.8%		
Television/Radio	8.1%	5.4%	6.6%	8.2%	7.9%	10.1%	11.6%	9.7%	9.8%		
Internet/Online services	44.4%	42.8%	49.0%	41.1%	44.9%	42.8%	39.8%	51.2%	48.6%		
Advertisements	5.6%	5.8%	3.8%	7.8%	5.8%	4.8%	9.0%	3.6%	2.2%		
Friends/Relatives	36.7%	44.9%	37.7%	39.3%	37.9%	33.4%	31.7%	27.8%	17.3%		
Lawyers	3.6%	3.3%	4.1%	3.7%	3.5%	2.7%	5.2%	1.3%	6.6%		
Accountants	14.0%	12.0%	15.5%	15.3%	10.7%	14.1%	15.1%	17.5%	19.6%		
Bankers	30.7%	37.1%	34.5%	37.2%	27.5%	24.5%	27.6%	21.4%	23.4%		
Brokers	17.3%	10.2%	19.2%	12.5%	18.8%	17.9%	22.2%	25.6%	27.0%		
Financial planners	42.4%	27.7%	41.5%	45.7%	36.8%	49.8%	54.8%	55.5%	51.0%		
Self	8.0%	7.9%	7.4%	8.4%	8.8%	7.5%	5.0%	9.3%	14.8%		
Financial planner, banker, broker, accountant, or lawyer	72.3%	65.5%	70.5%	73.4%	69.8%	77.7%	75.7%	80.8%	76.2%		

Source: 2013 SCF.

Note: Weighted by household weights. Percentages do not add to 100% because multiple answers were allowed.

Table 33. Sources of Financial Advice for Households with an IRA, by Household IRA Assets

Source	Households with IRAs		\$1- \$24,999		\$24,999 \$49,999		\$49,999 \$99,999	
	Sample size	2,170	571	788	1,095	14.6%	14.1%	7.3%
Call around		12.9%	15.9%	14.9%	14.6%			
Magazines/Newspapers/Books		18.5%	12.2%	13.9%	14.1%			
Material in the mail		7.2%	7.1%	8.1%	7.3%			
Television/Radio		8.1%	6.0%	6.5%	6.9%			
Internet/Online services		44.4%	45.8%	44.6%	44.7%			
Advertisements		5.6%	4.8%	5.6%	5.6%			
Friends/Relatives		36.7%	41.5%	40.9%	40.2%			
Lawyers		3.6%	3.7%	3.7%	3.6%			
Accountants		14.0%	13.7%	14.1%	13.2%			
Bankers		30.7%	35.9%	36.2%	34.0%			
Brokers		17.3%	14.5%	14.0%	15.2%			
Financial planners		42.4%	34.3%	37.2%	37.1%			
Self		8.0%	7.7%	7.8%	8.1%			
Financial planner, banker, broker, accountant, or lawyer		72.3%	67.9%	69.3%	69.4%			

Source: 2013 SCF.

Note: Weighted by household weights. Percentages do not add to 100% because multiple answers were allowed.

Table 34. IRA Assets of Households That Consulted a Financial Planner, Banker, Broker, Accountant, or Lawyer

Household IRA Assets	Households			IRA Assets		
	Number of households (millions)	Percent	Cumulative percent	Value of accounts (\$ billions)	Percent	Cumulative percent
None	44.3	64.0%	64.0%	0.0	0.0%	0.0%
\$1-\$9,999	4.2	6.1%	70.1%	17.5	0.4%	0.4%
\$10,000-\$24,999	4.2	6.1%	76.2%	67.0	1.4%	1.7%
\$25,000-\$49,999	3.1	4.5%	80.7%	109.7	2.2%	3.9%
\$50,000-\$99,999	3.9	5.7%	86.3%	275.0	5.5%	9.5%
\$100,000-\$249,999	4.6	6.6%	92.9%	709.3	14.3%	23.8%
\$250,000-\$499,999	2.4	3.5%	96.4%	845.5	17.0%	40.8%
\$500,000-\$999,999	1.5	2.2%	98.6%	1,017.6	20.5%	61.3%
\$1,000,000+	1.0	1.4%	100.0%	1,918.7	38.7%	100.0%
Total	69.2	100.0%	100.0%	4,960.3	100.0%	100.0%

Source: 2013 SCF.

Note: Weighted by household weights.

Table 35. IRA Assets of Households That Consulted a Broker

Household IRA Assets	Households			IRA Assets		
	Number of households (millions)	Percent	Cumulative percent	Value of accounts (\$ billions)	Percent	Cumulative percent
None	4.6	43.6%	43.6%	0.0	0.0%	0.0%
\$1 - \$10,000	0.7	6.2%	49.8%	3.3	0.2%	0.2%
\$10,000 - \$25,000	1.1	10.8%	60.6%	19.3	1.3%	1.6%
\$25,000 - \$50,000	0.5	5.0%	65.6%	19.4	1.3%	2.9%
\$50,000 - \$100,000	1.1	10.0%	75.6%	74.7	5.1%	8.0%
\$100,000 - \$250,000	1.1	9.9%	85.5%	162.7	11.2%	19.2%
\$250,000 - \$500,000	0.7	6.6%	92.1%	241.4	16.6%	35.7%
\$500,000 - \$1,000,000	0.5	4.6%	96.7%	329.7	22.6%	58.3%
\$1,000,000 and above	0.3	3.3%	100.0%	607.9	41.7%	100.0%
Total	10.6	100.0%		1,458.4	100.0%	

Source: 2013 SCF.

Note: Weighted by household weights.

Table 36. Median Household Income, by IRA Ownership and Professional Advice

	Number of households (millions)	Median household income (\$ thousands)	IRA-asset-weighted median household income (\$ thousands)
No IRA, no professional advice	43.8	35.5	
No IRA, receives professional advice	44.3	45.7	
Owens IRA, no professional advice	9.5	86.2	182.6
Owens IRA, receives professional advice	24.9	96.3	166.4
All	122.5	50.7	175.5

Source: 2013 SCF.

Note: Weighted by household weights. Professional advice includes advice from a financial planner, banker, broker, accountant, or lawyer.

Table 37. Sources of Financial Advice by Household Income

Source	Annual household income													
	Under \$10,000		\$10,000 - \$20,000		\$20,000 - \$50,000		\$50,000 - \$100,000		\$100,000 - \$200,000		\$200,000 - \$500,000		\$500,000 and above	
	Sample size	146	586	1,673	1,479	952	462	718						
Call around	15.7%	12.6%	13.4%	13.5%	13.1%	10.2%	10.6%							
Magazines/Newspapers/Books	6.4%	6.6%	8.8%	10.4%	14.4%	26.3%	29.4%							
Material in the mail	3.8%	6.2%	6.1%	6.6%	7.2%	6.1%	5.4%							
Television/Radio	6.0%	8.0%	7.5%	8.4%	8.5%	8.1%	8.0%							
Internet/Online services	23.2%	18.1%	27.9%	38.1%	52.1%	60.3%	53.0%							
Advertisements	5.8%	5.7%	6.7%	6.4%	5.6%	5.1%	3.2%							
Friends/Relatives	41.9%	38.0%	40.3%	40.7%	40.4%	39.2%	34.2%							
Lawyers	6.4%	4.2%	3.1%	3.6%	4.1%	5.9%	7.0%							
Accountants	4.3%	8.0%	7.2%	11.0%	14.1%	17.7%	25.5%							
Bankers	24.9%	31.9%	34.3%	35.7%	29.9%	25.4%	32.5%							
Brokers	2.8%	3.1%	5.5%	8.8%	13.4%	24.3%	32.4%							
Financial planners	15.2%	14.2%	19.0%	28.5%	37.6%	40.5%	49.3%							
Self	2.8%	4.4%	5.1%	7.3%	9.1%	13.8%	14.0%							

Source: 2013 SCF.

Note: Weighted by household weights. Percentages do not add to 100% because multiple answers were allowed.

Table 38. Sources of Financial Advice by Head of Household Demographics

Source	Ethnicity		Race						Age					
	Hispanic /Latino origin	Non-Hispanic /Latino	White	Black	Hispanic	Other	18-25	26-35	36-45	46-55	56-65	66-75	76-85	>85
	712	5,303	4,425	746	556	288	288	828	1,065	1,342	1,259	759	361	113
Sample size														
Call around	12.6%	13.3%	12.6%	15.8%	12.7%	14.0%	15.6%	14.5%	11.5%	15.3%	13.7%	10.5%	11.8%	7.1%
Magazines/Newspapers/Books	8.8%	11.3%	11.3%	9.6%	8.8%	15.4%	6.2%	8.3%	9.7%	11.1%	13.0%	16.1%	10.6%	8.8%
Material in the mail	7.1%	6.3%	5.9%	7.5%	7.7%	6.9%	3.2%	6.2%	6.1%	6.8%	7.2%	7.9%	5.1%	3.4%
Television/Radio	10.6%	7.5%	7.0%	9.6%	11.2%	10.2%	6.4%	5.0%	6.2%	10.8%	9.3%	10.5%	6.0%	4.1%
Internet/Online services	29.9%	36.1%	36.4%	31.0%	27.1%	50.0%	42.9%	47.5%	40.3%	40.0%	33.0%	23.8%	11.1%	7.4%
Advertisements	6.6%	6.1%	5.6%	7.7%	6.7%	8.3%	7.8%	6.8%	5.5%	7.1%	6.6%	4.9%	4.8%	1.9%
Friends/Relatives	37.9%	40.4%	40.1%	39.7%	37.1%	46.5%	63.4%	52.3%	40.7%	39.8%	34.5%	28.3%	29.7%	29.9%
Lawyers	4.1%	3.8%	3.7%	3.9%	3.6%	5.2%	3.6%	3.1%	3.7%	4.3%	4.5%	3.4%	3.7%	4.7%
Accountants	6.6%	10.7%	10.8%	9.7%	5.6%	12.8%	9.2%	9.4%	9.4%	12.7%	10.0%	11.3%	8.0%	5.5%
Bankers	35.5%	32.6%	32.3%	33.2%	34.2%	39.9%	32.7%	36.2%	31.3%	33.8%	33.6%	30.3%	31.5%	32.0%
Brokers	6.2%	9.0%	9.6%	5.3%	5.7%	10.7%	2.2%	5.4%	7.3%	10.2%	11.0%	12.9%	6.8%	9.1%
Financial planners	17.0%	26.8%	28.0%	22.6%	13.7%	22.6%	15.6%	22.8%	23.9%	28.0%	31.8%	27.7%	20.2%	15.9%
Self	4.2%	7.2%	7.2%	7.2%	4.1%	5.4%	3.6%	7.2%	6.4%	5.6%	7.9%	7.8%	7.1%	10.3%

Source: 2013 SCF.

Note: Weighted by household weights. Percentages do not add to 100% because multiple answers were allowed.

Table 39. Sources of Financial Advice by Financial Institution(s) that Hold Households' IRAs

Source	Finance or				Investment			
	Commercial bank	Savings and loan	Credit union	Finance or loan company	Insurance company	Brokerage company	Investment /mgmt company	Other
Sample size	1,086	60	129	33	43	20	29	
Call around	11.9%	16.8%	10.1%	16.1%	20.4%	2.6%	9.3%	
Magazines/Newspapers/Books	18.6%	24.6%	11.0%	8.7%	26.4%	19.3%	14.7%	
Material in the mail	6.9%	9.3%	1.5%	12.4%	11.4%	14.2%	1.7%	
Television/Radio	6.8%	14.9%	8.5%	11.8%	9.7%	14.2%	6.6%	
Internet/Online services	42.5%	37.8%	35.4%	40.1%	54.4%	42.2%	25.8%	
Advertisements	5.6%	10.1%	7.1%	8.5%	15.1%	16.8%	7.5%	
Friends/Relatives	34.3%	32.9%	31.2%	32.2%	34.4%	18.1%	42.2%	
Lawyers	3.1%	0.0%	2.3%	4.7%	7.2%	10.4%	0.0%	
Accountants	11.9%	17.6%	9.7%	4.8%	10.7%	21.0%	15.1%	
Bankers	32.4%	53.0%	41.9%	21.0%	36.3%	18.4%	24.3%	
Brokers	14.6%	18.8%	11.1%	12.2%	23.9%	6.6%	6.1%	
Financial planners	43.7%	38.4%	26.7%	43.7%	50.2%	62.9%	57.5%	
Self	9.1%	2.1%	15.8%	12.1%	14.9%	0.0%	3.7%	

Source: 2013 SCF.

Note: Weighted by household weights. Percentages do not add to 100% because multiple answers were allowed.

Table 40. Households' IRA-holding Financial Institution(s) by Total Household IRA Assets

IRA-holding Institution	All households with IRAs	Household IRA Assets									
		\$1- \$9,999	\$10,000- \$24,999	\$25,000- \$49,999	\$50,000- \$99,999	\$100,000- \$249,999	\$250,000- \$499,999	\$500,000- \$999,999	\$1,000,000- and above	Sample size	
	2,170	296	275	217	308	382	251	205	237		
Commercial bank	49.7%	46.7%	47.6%	54.1%	50.0%	50.8%	50.0%	49.7%	51.6%		
Savings and loan	4.1%	2.4%	5.5%	5.0%	6.6%	3.7%	2.7%	1.3%	0.0%		
Credit union	8.8%	16.1%	10.6%	10.2%	8.0%	6.4%	2.5%	0.6%	2.8%		
Finance or loan company	1.5%	1.1%	1.0%	2.4%	1.2%	1.9%	0.8%	2.8%	2.1%		
Brokerage	41.2%	31.6%	36.7%	34.7%	39.4%	50.3%	50.8%	55.8%	52.2%		
Insurance company	2.2%	1.6%	2.0%	2.4%	1.5%	2.3%	4.7%	3.6%	0.8%		
Investment/mgmt company	0.7%	0.9%	0.0%	0.7%	0.8%	0.3%	1.5%	0.6%	3.1%		
Other	1.1%	0.7%	0.4%	0.8%	1.6%	1.6%	2.3%	0.3%	1.8%		

Source: 2013 SCF.

Note: Weighted by household weights. Percentages do not add to 100% because multiple answers were allowed.

Table 41. Households' IRA-holding Financial Institution(s) by Total Household Income

IRA-holding Institution	Annual Household Income															
	All Households with IRA		Under \$10,000		\$10,000 - \$20,000		\$20,000 - \$50,000		\$50,000 - \$100,000		\$100,000 - \$200,000		\$200,000 - \$500,000		\$500,000 and above	
	Sample size	2,170	8	18	257	498	529	336	523							
Commercial bank	49.7%	57.1%	51.9%	49.9%	51.7%	48.7%	45.9%	49.8%								
Savings and loan	4.1%	0.0%	7.1%	5.8%	4.5%	3.9%	1.3%	0.7%								
Credit union	8.8%	36.8%	12.1%	11.2%	12.0%	6.0%	5.2%	0.5%								
Finance or loan company	1.5%	0.0%	0.0%	0.9%	1.6%	2.2%	0.4%	2.3%								
Brokerage	41.2%	0.2%	20.7%	34.2%	36.3%	46.1%	52.3%	54.4%								
Insurance company	2.2%	5.9%	0.0%	2.5%	2.2%	1.7%	3.7%	1.9%								
Investment/mgmt company	0.7%	0.0%	2.5%	1.0%	0.6%	0.4%	1.3%	0.9%								
Other	1.1%	0.0%	5.7%	0.8%	1.6%	0.8%	1.2%	1.1%								

Source: 2013 SCF.

Note: Weighted by household weights. Percentages do not add to 100% because multiple answers were allowed.

Table 42. Households' IRA-holding Financial Institution(s) by Head of Household Demographics

IRA-holding Institution	Marital Status						Education						
	Married	Separated	Divorced	Widowed	Never married- Male	Never married- Female	Less than high school	High school or	Some college	Assoc degree	Bachelor degree	Master degree	Prof. degree/ doctorat
Sample size	3,286	201	920	446	686	476	544	1,601	1,028	369	1,346	726	400
Commercial bank	50.6%	29.6%	45.4%	53.2%	48.5%	49.4%	39.8%	50.9%	49.3%	55.7%	51.0%	45.5%	48.1%
Savings and loan	3.9%	0.0%	4.8%	6.4%	3.6%	2.7%	0.0%	5.8%	4.9%	5.3%	3.1%	3.3%	3.6%
Credit union	7.5%	30.6%	14.1%	4.2%	12.6%	9.5%	17.8%	15.1%	10.5%	7.2%	6.4%	6.8%	4.4%
Finance or loan company	1.7%	0.0%	1.3%	0.8%	1.1%	2.1%	0.0%	2.2%	0.6%	1.5%	0.8%	3.3%	0.3%
Brokerage	43.8%	40.0%	35.9%	39.2%	36.5%	32.1%	47.1%	29.3%	40.3%	30.9%	45.2%	46.2%	52.9%
Insurance company	2.2%	0.0%	2.7%	4.1%	1.5%	0.5%	0.0%	2.2%	2.0%	3.3%	1.8%	2.8%	2.5%
Investment/mgmt company	0.8%	0.0%	0.1%	1.1%	0.3%	1.5%	0.0%	0.5%	0.3%	1.3%	0.5%	1.3%	1.5%
Other	1.2%	0.0%	1.6%	0.7%	0.0%	2.6%	0.0%	0.9%	0.4%	1.1%	2.1%	1.0%	0.2%

IRA-holding Institution	Ethnicity			Race			Age							
	Hispanic /Latino origin	Non-Hispanic /Latino	Other	White	Black	Hispanic	18-25	26-35	36-45	46-55	56-65	66-75	76-85	>85
Sample size	712	5,303	4,425	746	556	288	288	828	1,065	1,342	1,259	759	361	113
Commercial bank	50.3%	49.7%	50.2%	38.7%	53.6%	51.4%	52.3%	44.2%	49.2%	47.5%	51.3%	48.8%	56.0%	69.9%
Savings and loan	3.9%	4.1%	3.9%	7.6%	0.0%	5.2%	0.0%	3.5%	2.4%	3.4%	3.2%	6.4%	7.9%	4.1%
Credit union	7.1%	8.9%	8.5%	19.6%	8.6%	1.9%	5.1%	11.6%	7.6%	8.9%	9.9%	7.4%	8.0%	6.2%
Finance or loan company	1.6%	1.5%	1.3%	2.7%	2.8%	3.0%	0.0%	2.4%	1.1%	3.4%	0.9%	0.7%	0.3%	0.0%
Brokerage	36.4%	41.4%	42.0%	31.4%	31.1%	43.3%	30.0%	38.4%	43.7%	45.8%	39.6%	42.7%	33.1%	28.8%
Insurance company	1.3%	2.3%	2.4%	2.2%	2.3%	0.3%	0.0%	1.0%	1.2%	1.2%	3.4%	3.5%	2.6%	1.1%
Investment/mgmt company	0.1%	0.8%	0.7%	0.9%	0.0%	1.9%	12.6%	1.0%	0.5%	0.4%	0.6%	1.3%	0.2%	0.0%
Other	1.4%	1.2%	1.1%	1.1%	2.5%	1.8%	0.0%	0.6%	1.0%	1.1%	1.2%	2.0%	1.0%	0.0%

Source: 2013 SCF.

Note: Weighted by household weights. Percentages do not add to 100% because multiple answers were allowed.

Table 43. Households' IRA-holding Financial Institution(s) by Total Household Wealth Quintiles and Top Percentiles

IRA-holding Institution	Total household wealth quintiles					Top 5%	Top 1%
	Lowest- First	Second	Third	Fourth	Highest- Fifth		
Sample size	1,242	787	986	1,015	1,984	995	510
Lower wealth boundary	\$0	\$0	\$3,300	\$30,000	\$165,000	\$881,000	\$3,700,000
Commercial bank	0.0%	45.3%	46.8%	51.1%	49.8%	48.7%	51.6%
Savings and loan	0.0%	3.1%	2.7%	6.2%	3.2%	1.0%	0.8%
Credit union	0.0%	13.7%	14.6%	12.9%	4.7%	3.2%	0.5%
Finance or loan company	0.0%	2.4%	1.3%	1.3%	1.6%	1.2%	1.7%
Brokerage	0.0%	25.4%	33.6%	33.5%	48.5%	53.6%	56.1%
Insurance company	0.0%	2.9%	2.2%	1.2%	2.8%	3.3%	1.5%
Investment/mgmt company	0.0%	3.3%	0.5%	0.7%	0.7%	1.1%	1.9%
Other	0.0%	3.9%	0.6%	0.4%	1.6%	0.7%	0.2%

Source: 2013 SCF.

Note: Weighted by household weights. Percentages do not add to 100% because multiple answers were allowed.

Table 44. Distribution of Asset Holdings by Total Household Wealth Quintiles

Account Type	Total household wealth quintiles (lower wealth boundary)											
	Lowest-First (\$0)		Second (\$0)		Third (\$3,300)							
	\$ billions	% of Assets	% of HH with positive balance	% of HH among HH with positive balance	\$ billions	% of Assets	% of HH with positive balance	% of HH among HH with positive balance	\$ billions	% of Assets	% of HH with positive balance	% of HH among HH with positive balance
DC accounts	0.0	NA	NA	NA	3.4	17.7%	14.9%	78.4%	107.6	32.9%	42.1%	72.8%
IRAs	0.0	NA	NA	NA	1.3	6.9%	5.1%	82.7%	44.6	13.6%	18.9%	60.9%
Stocks	0.0	NA	NA	NA	0.4	2.1%	2.9%	61.6%	13.2	4.0%	7.7%	42.8%
Mutual funds	0.0	NA	NA	NA	0.2	1.2%	1.2%	59.9%	4.0	1.2%	2.3%	49.9%
Savings accounts	0.0	NA	NA	NA	9.8	50.8%	73.6%	73.3%	94.8	29.0%	64.1%	44.6%
Bonds	0.0	NA	NA	NA	0.0	0.0%	0.0%	0.0%	0.3	0.1%	0.1%	81.6%
CDs	0.0	NA	NA	NA	0.3	1.5%	1.3%	73.8%	12.7	3.9%	7.5%	46.6%
Trusts	0.0	NA	NA	NA	0.0	0.0%	0.0%	0.0%	0.6	0.2%	0.2%	53.1%
Cashable life	0.0	NA	NA	NA	3.0	15.5%	15.1%	79.5%	41.2	12.6%	24.3%	53.5%
Cashable annuities	0.0	NA	NA	NA	0.0	0.0%	0.1%	16.2%	4.3	1.3%	1.5%	65.0%
Broker cash	0.0	NA	NA	NA	0.0	0.0%	0.0%	0.0%	0.0	0.0%	0.1%	4.2%
Savings bonds	0.0	NA	NA	NA	0.8	4.3%	6.6%	59.4%	3.8	1.2%	9.4%	12.2%
Total	0.0	NA	NA	NA	19.4	100.0%	6.6%	59.4%	327.0	100.0%	9.4%	12.2%

(Continued on next page)

Table 44. Distribution of Asset Holdings by Total Household Wealth Quintiles (Continued)

Account Type	Total household wealth quintiles (lower wealth boundary)					% of HH with positive balance	% of assets among HH with positive balance	% of HH with positive balance	% of HH among HH with positive balance	% of assets among HH with positive balance
	\$ billions	% of Assets	% of Assets with positive balance	\$ billions	% of Assets					
DC accounts	753.5	38.1%	58.4%	4,443.4	16.9%	58.1%	33.3%			
IRAs	465.7	23.5%	43.9%	6,164.3	23.5%	73.9%	28.8%			
Stocks	93.7	4.7%	16.7%	4,685.1	17.9%	42.3%	28.3%			
Mutual funds	62.0	3.1%	8.6%	4,482.4	17.1%	29.1%	33.5%			
Savings accounts	318.9	16.1%	72.7%	1,553.4	5.9%	72.1%	9.2%			
Bonds	4.7	0.2%	0.8%	954.4	3.6%	6.3%	16.6%			
CDs	91.1	4.6%	12.8%	507.9	1.9%	17.5%	12.4%			
Trusts	11.8	0.6%	1.1%	1,675.2	6.4%	6.1%	42.0%			
Cashable life	99.0	5.0%	26.4%	685.4	2.6%	33.6%	6.9%			
Cashable annuities	59.0	3.0%	5.2%	609.5	2.3%	12.6%	17.9%			
Broker cash	4.3	0.2%	1.4%	405.4	1.5%	7.8%	9.8%			
Savings bonds	16.0	0.8%	15.5%	58.0	0.2%	20.1%	1.2%			
Total	1,979.7	100.0%	100.0%	26,224.4	100.0%					

Source: 2013 SCF.

Note: Weighted by household weights. Total household wealth is the sum of DC accounts, IRAs, stocks, mutual funds, savings accounts, bonds, CDs, trusts, cashable life, cashable annuities, broker cash, and savings bonds.

Table 45. Distribution of Asset Holdings Overall and by Total Household Wealth Top Percentiles

Account Type	Total household wealth percentiles (lower wealth boundary)														
	All Households					Top 5% (\$881,000)					Top 1% (\$3,700,000)				
	\$ billions	% of Assets	% of HH with positive balance	% of assets among HH with positive balance	\$ billions	% of Assets	% of HH with positive balance	% of assets among HH with positive balance	\$ billions	% of Assets	% of HH with positive balance	% of assets among HH with positive balance			
DC accounts	5,308.0	18.6%	34.0%	36.2%	2,088.3	10.9%	53.1%	23.0%	824.1	7.5%	51.3%	16.8%			
IRAs	6,676.0	23.4%	28.1%	29.8%	4,122.1	21.6%	86.8%	25.3%	1,785.9	16.2%	84.5%	19.6%			
Stocks	4,792.3	16.8%	13.8%	28.4%	4,060.3	21.3%	62.8%	29.6%	2,611.6	23.7%	79.5%	30.0%			
Mutual funds	4,548.7	15.9%	8.2%	33.5%	3,944.6	20.7%	50.5%	33.8%	2,643.9	24.0%	61.1%	35.7%			
Savings accounts	1,977.0	6.9%	53.2%	10.7%	976.4	5.1%	68.0%	8.4%	436.4	4.0%	60.1%	7.3%			
Bonds	959.3	3.4%	1.4%	16.6%	927.1	4.9%	17.8%	16.6%	723.1	6.6%	37.2%	16.6%			
CDs	612.0	2.1%	7.8%	13.9%	287.2	1.5%	16.6%	10.2%	111.7	1.0%	15.7%	7.4%			
Trusts	1,687.6	5.9%	1.5%	42.0%	1,502.1	7.9%	11.4%	41.6%	1,214.6	11.0%	20.1%	41.8%			
Cashable life	828.7	2.9%	19.2%	7.9%	460.4	2.4%	39.1%	6.1%	275.5	2.5%	38.2%	6.3%			
Cashable annuities	672.7	2.4%	3.9%	19.0%	332.8	1.7%	13.8%	13.4%	111.3	1.0%	12.6%	8.4%			
Broker cash	409.8	1.4%	1.9%	9.8%	376.5	2.0%	14.7%	10.3%	256.2	2.3%	20.8%	9.9%			
Savings bonds	78.6	0.3%	10.0%	1.5%	22.5	0.1%	19.1%	0.7%	8.3	0.1%	18.4%	0.5%			
Total	28,550.5	100.0%			19,100.3	100.0%			11,002.8	100.0%					

Source: 2013 SCF.

Note: Weighted by household weights. Total household wealth is the sum of DC accounts, IRAs, stocks, mutual funds, savings accounts, bonds, CDs, trusts, cashable life, cashable annuities, broker cash, and savings bonds.

Table 46. Distribution of Asset Holdings among IRA Holders by Total IRA-holding Household Wealth Quintiles

Account Type	Total household wealth quintiles (lower wealth boundary)											
	Lowest-First (\$0)		Second (\$40,170)		Third (\$121,800)		Fourth (\$270,000)		Highest-Fifth (\$650,000)			
	\$ billions	% of Assets	% of HH with positive balance	% of HH among assets with HH with positive balance	\$ billions	% of Assets	% of HH with positive balance	% of HH among assets with HH with positive balance	\$ billions	% of Assets	% of HH with positive balance	% of HH among assets with HH with positive balance
DC accounts	16.3	13.6%	26.2%	41.8%	118.3	21.8%	47.2%	45.6%	327.2	25.3%	52.2%	48.7%
IRAs	69.3	58.0%	100.0%	58.0%	273.4	50.4%	100.0%	50.4%	559.4	43.2%	100.0%	43.2%
Stocks	2.0	1.7%	8.9%	16.5%	16.6	3.1%	16.5%	18.0%	57.3	4.4%	27.1%	15.7%
Mutual funds	1.9	1.6%	3.9%	28.6%	17.3	3.2%	10.6%	26.6%	59.5	4.6%	17.2%	25.7%
Savings accounts	19.8	16.5%	55.4%	27.5%	67.9	12.5%	70.6%	17.8%	144.0	11.1%	75.2%	14.7%
Bonds	0.0	0.0%	0.0%	0.0%	0.1	0.0%	0.0%	50.7%	3.3	0.3%	2.5%	8.3%
CDs	3.4	2.8%	6.6%	34.7%	15.4	2.8%	11.6%	24.8%	31.0	2.4%	16.5%	14.7%
Trusts	0.0	0.0%	0.0%	0.0%	0.9	0.2%	0.5%	40.1%	3.5	0.3%	0.8%	40.4%
Cashable life	5.0	4.2%	14.5%	25.9%	19.2	3.5%	27.6%	12.1%	53.2	4.1%	29.1%	14.1%
Cashable annuities	0.1	0.1%	1.3%	9.9%	8.7	1.6%	4.3%	30.1%	43.4	3.4%	9.6%	33.4%
Broker cash	0.0	0.0%	0.0%	0.0%	2.2	0.4%	3.0%	12.3%	3.3	0.3%	3.3%	6.4%
Savings bonds	1.7	1.4%	13.1%	9.6%	3.0	0.6%	14.5%	3.8%	9.3	0.7%	22.0%	3.3%
Total	119.5	100.0%			542.9	100.0%			1,294.3	100.0%		
									2,998.6	100.0%		
									17,473.2	100.0%		

Source: 2013 SCF.

Note: Weighted by household weights. Total household wealth is the sum of DC accounts, IRAs, stocks, mutual funds, savings accounts, bonds, CDs, trusts, cashable life, cashable annuities, broker cash, and savings bonds.

Table 47. Distribution of Asset Holdings among IRA Holders Overall and by Total IRA-holding Household Wealth Top Percentiles

Account Type	All Households with IRA										Total household wealth percentiles (lower wealth boundary)										
	%					%					%					%					
	\$ billions	% of Assets	% of HH with positive balance	% of HH among HH with positive balance	% of assets among HH with positive balance	\$ billions	% of Assets	% of HH with positive balance	% of HH among HH with positive balance	% of assets among HH with positive balance	\$ billions	% of Assets	% of HH with positive balance	% of HH among HH with positive balance	% of assets among HH with positive balance	\$ billions	% of Assets	% of HH with positive balance	% of HH among HH with positive balance	% of assets among HH with positive balance	
DC accounts	3,041.1	13.6%	46.9%	28.1%	728.4	6.5%	49.0%	14.8%	164.6	3.0%	40.6%	8.0%	22,428.5	100%	11,233.5	100%	5,414.6	100%			
IRAs	6,676.0	29.8%	100.0%	29.8%	2,336.6	20.8%	100.0%	20.8%	939.5	17.4%	100.0%	17.4%	1,188.2	5.3%	90.0	0.8%	12.1%	13.6	0.3%	14.4%	2.0%
Stocks	3,771.2	16.8%	30.4%	26.7%	2,655.1	23.6%	80.2%	29.4%	1,292.3	23.9%	76.2%	31.1%	574.8	2.6%	971.6	8.6%	14.0%	684.0	12.6%	29.9%	35.8%
Mutual funds	3,715.9	16.6%	20.7%	32.0%	2,540.1	22.6%	68.8%	32.1%	1,463.9	27.0%	72.8%	35.4%	511.6	2.3%	273.5	2.4%	42.1%	140.6	2.6%	38.6%	6.5%
Savings accounts	1,345.5	6.0%	68.8%	9.3%	502.2	4.5%	63.4%	7.8%	109.3	2.0%	55.3%	4.0%	341.7	1.5%	134.9	1.2%	15.2%	72.9	1.3%	13.4%	10.4%
Bonds	823.5	3.7%	4.0%	16.1%	697.2	6.2%	34.4%	15.9%	326.9	6.0%	40.0%	14.2%	54.8	0.2%	291.4	2.6%	20.8%	205.3	3.8%	33.2%	12.3%
CDs	384.2	1.7%	14.4%	11.0%	90.0	0.8%	12.1%	6.4%	13.6	0.3%	14.4%	2.0%	54.8	0.2%	12.5	0.1%	21.2%	1.6	0.0%	19.0%	0.2%
Trusts	1,188.2	5.3%	3.0%	36.6%	90.0	0.8%	12.1%	6.4%	13.6	0.3%	14.4%	2.0%	54.8	0.2%	12.5	0.1%	21.2%	1.6	0.0%	19.0%	0.2%
Cashable life	574.8	2.6%	28.8%	6.7%	273.5	2.4%	42.1%	5.9%	140.6	2.6%	38.6%	6.5%	511.6	2.3%	134.9	1.2%	15.2%	72.9	1.3%	13.4%	10.4%
Cashable annuities	511.6	2.3%	8.8%	16.4%	134.9	1.2%	15.2%	8.4%	72.9	1.3%	13.4%	10.4%	341.7	1.5%	291.4	2.6%	20.8%	205.3	3.8%	33.2%	12.3%
Broker cash	341.7	1.5%	5.4%	9.1%	291.4	2.6%	20.8%	10.7%	205.3	3.8%	33.2%	12.3%	54.8	0.2%	12.5	0.1%	21.2%	1.6	0.0%	19.0%	0.2%
Savings bonds	54.8	0.2%	18.0%	1.3%	12.5	0.1%	21.2%	0.6%	1.6	0.0%	19.0%	0.2%	22,428.5	100%	11,233.5	100%	5,414.6	100%			

Source: 2013 SCF.

Note: Weighted by household weights. Total household wealth is the sum of DC accounts, IRAs, stocks, mutual funds, savings accounts, bonds, CDs, trusts, cashable life, cashable annuities, broker cash, and savings bonds.

APPENDIX 1

This section's facts, statements, and tabulations are current as of September 30, 2011.

Availability of Data on IRA and DC Account Age

We researched the availability of information on the duration from the purchase to the sale of assets, but were unable to locate such information. However, the SIPP and SCF do ask some questions about the age of IRA and DC plan accounts.

SIPP

The SIPP Assets and Liabilities topical module contains the following questions on IRA and DC account age.

- For how many years has ... contributed to ...'s IRA accounts?
- For how many years has ... contributed to ...'s 401k, 403b, or thrift plans?

The SIPP Retirement and Pension Plans topical module contains the following question for DC accounts.

- How many years have you been included in this plan?

SCF

The SCF does not include any data on IRA account age, but it includes the following question on the DC pension plan with their current main job.

- How long (have you/has your[husband/wife/partner]) been in this plan?
 - Number of years
 - Age of respondent when account opened
 - Year plan began

"How America Saves 2011" Vanguard

A thorough review of Vanguard's "How America Saves" report revealed that Vanguard does not publish information on the duration that assets are held or the age of accounts.⁶

EBRI

A cursory review of EBRI publications and the EBRI website did not reveal any information on the duration that assets are held or the age of IRAs or DC plans.

⁶ Downloaded from <https://institutional.vanguard.com/iam/pdf/HAS11.pdf>.

APPENDIX 2

Calculations for Table 26: IRAs, IRA assets, and Median Household Income of IRA Holders, by Type of Institution Holding IRAs

The SCF does not ask respondents to specify the type of institution in which each IRA account is held. We used the responses to two sets of questions to assign accounts and balances to institutions. First, the SCF asks about total balances in four different types of IRA accounts: Roth, roll-over from pension account, regular or other IRA, and Keogh. These data were collected separately for the respondent, spouse and other family members who are part of the primary economic unit, if any. Second, respondents were asked to list the type of each institution (e.g., commercial bank, brokerage) where any IRA accounts were held. Again, this was asked separately for the respondent, spouse and other family members and responses were recorded in the order that they were given.

Each positive balance reported in the first set of questions was treated as if it related to one account. Accounts and balances were assigned to the institution types reported in the following manner:

- 1) If only one institution type was reported, all accounts were assumed to be held at that institution type;
- 2) If multiple types of institutions were reported, the account balances were sorted and the account with the largest balance was assigned to the institution mentioned first, the account with the second largest balance was assigned to the institution mentioned second and so on;
- 3) If the number of accounts reported was larger than the number of institutions reported then the account with the smallest balance was assigned to the last mentioned institution, the next smallest to the next-to-last mentioned institution and so on, with the remaining balances all assigned to the first mentioned institution;
- 4) If the number of balances was smaller than the number of institutions then we used the process described in (2) and ignored any excess institutions.

DISCLAIMER

The views, opinions, and/or findings contained in this report should not be construed as an official Government position, policy or decision, unless so designated by other documentation issued by the appropriate governmental authority.

We call your attention to the possibility that other professionals may perform procedures concerning the same information or data and reach different findings than Advanced Analytical Consulting Group, Inc. (AACG) and Deloitte Transaction and Business Analytics LLP (Deloitte) for a variety of reasons, including the possibilities that additional or different information or data might be provided to them that was not provided to Deloitte and AACG, that they might perform different procedures than did Deloitte and AACG, or that professional judgments concerning complex, unusual, or poorly documented matters may differ.

This document contains general information only. AACG and Deloitte are not, by means of this document, rendering business, financial, investment, or other professional advice or services. This document is not a substitute for such professional advice or services, nor should it be used as a basis for any decision or action. Before making any decision or taking any action, a qualified professional adviser should be consulted. AACG and Deloitte, its affiliates, or related entities shall not be responsible for any loss sustained by any person who relies on this publication.

84

NBER WORKING PAPER SERIES

IT PAYS TO SET THE MENU:
MUTUAL FUND INVESTMENT OPTIONS IN 401(K) PLANS

Veronika K. Pool
Clemens Sialm
Irina Stefanescu

Working Paper 18764
<http://www.nber.org/papers/w18764>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2013

We thank Pierluigi Balduzzi, Keith Brown, Lauren Cohen, Van Harlow, Frank de Jong, Olivia Mitchell, Jonathan Reuter, Paul Schultz, Laura Starks, Steve Utkus, Marno Verbeek, Scott Yonker, and seminar participants at the American Economic Association Meeting in San Diego, DePaul University, Indiana University, INSEAD, the IU-Notre Dame-Purdue Summer Symposium, the NETSPAR spring workshop, the Nova Finance Conference on Pensions and Retirement, and the Second MSUFCU Conference on Financial Institutions and Investments at Michigan State University for helpful comments, and NETSPAR, Indiana University, and the University of Texas at Austin for financial support. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2013 by Veronika K. Pool, Clemens Sialm, and Irina Stefanescu. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

It Pays to Set the Menu: Mutual Fund Investment Options in 401(k) Plans
Veronika K. Pool, Clemens Sialm, and Irina Stefanescu
NBER Working Paper No. 18764
February 2013
JEL No. G02,G11,G23,H55

ABSTRACT

This paper investigates whether mutual fund families acting as trustees of 401(k) plans display favoritism toward their own funds. Using a hand-collected dataset on retirement investment options, we show that poorly-performing funds are less likely to be removed from and more likely to be added to a 401(k) menu if they are affiliated with the plan trustee. We find no evidence that plan participants undo this affiliation bias through their investment choices. Finally, the subsequent performance of poorly-performing affiliated funds indicates that these trustee decisions are not information driven and are costly to retirement savers.

Veronika K. Pool
Kelley School of Business
Indiana University
1309 E 10th Street
Bloomington, IN 47405
vkpool@indiana.edu

Irina Stefanescu
Kelley School of Business
Indiana University
1309 E. Tenth Street
Bloomington, IN 47405
stefanei@indiana.edu

Clemens Sialm
University of Texas at Austin
McCombs School of Business
1 University Station; B6600
Austin, TX 78712
and NBER
clemens.sialm@mcombs.utexas.edu

1 Introduction

Employer-sponsored 401(k) accounts have gained significant importance around the world. In the United States, the value of 401(k) assets reached \$3.5 trillion at the end of the third quarter of 2012.¹ Their growth represents important business opportunities for mutual funds as they manage approximately half of the 401(k) investment pool. Moreover, mutual fund families often serve as trustees of these defined contribution (DC) plans and play an active role in creating the menu of investment options for the plans' participants.²

While the Employee Retirement Income Security Act of 1974 (ERISA) requires trustees to be prudent in selecting a suitable set of investment choices for their 401(k) clients, mutual fund trustees have a competing interest to maximize investments in their own proprietary funds. Surprisingly, little is known about whether and how these conflicting incentives influence the menu of options in 401(k) plans. This is concerning given that DC accounts are a main source of retirement income for many of the beneficiaries. Since retirement savings compound over long horizons, any inefficiency or trustee bias in this setting can significantly affect the employees' wealth at retirement and thus have important welfare consequences for society in general.

In this paper, we provide the first study of the conflicting incentives of mutual fund trustees. Focusing on fund entry and exit, we hypothesize that if the trustees' decisions are driven by their own financial interests, mutual fund trustees may be more inclined to include their own funds in the fund lineup – even when more suitable options are available from other fund families – and subsequently more reluctant to remove them. Moreover, they may also be less

¹Federal Reserve Statistical Releases and Investment Company Institute (ICI).

²The trustee is the entity who holds the assets of the plan in trust. Trustees are typically appointed by the employer who sponsors the plan and have a fiduciary responsibility to administer the plan for the exclusive benefit of plan members. They must act in accordance with reasonable standards of prudence and offer a diversified set of options to participants and sponsors. Cohen and Schmidt (2009) describe in detail the role of the trustee in the decision process.

sensitive to the performance of these affiliated funds when making menu altering decisions.

To investigate this *favoritism* hypothesis, we hand collect information on the menu of mutual fund options offered in a large sample of defined contribution plans for the period 1998 to 2009 from annual filings of Form 11-K with the U.S. Securities and Exchange Commission (SEC). Our sample includes plans that are trustee by a mutual fund family as well as plans with non-mutual fund trustees. Most 401(k) plans in our sample that have a mutual fund trustee adopt an *open architecture* whereby investment options include not only funds from the trustee's family but from other mutual fund families as well. Therefore, an interesting feature of our dataset is that a given fund often contemporaneously appears on several 401(k) menus that are administered by different trustees. This means that open architecture allows the same fund to be an affiliated fund ("trustee fund") on some menus and an unaffiliated fund ("non-trustee fund") on others. This data feature provides us with a unique identification strategy and allows us to contrast how the very same fund is viewed across two different menus: one on which it is a trustee fund and another on which it is not.

We find that despite their fiduciary responsibilities, trustees have a strong preference for their own funds. Trustee funds are less likely to be removed from the plan across the board. Moreover, the biggest difference between how trustee and non-trustee funds are treated on the menu occurs for the worst performing funds, which have been shown to exhibit significant performance persistence (Carhart, 1997). For example, mutual funds ranked in the lowest decile based on past performance (among the universe of funds in the same style category over the prior 36 months), are approximately two and a half times more likely to be deleted from those menus on which they are unaffiliated with the trustee than from those where they are affiliated with the trustee.

Similarly, we find that trustees are substantially more likely to add their own funds to the menu across all performance deciles. Furthermore, trustee fund additions exhibit lower

prior performance than non-trustee additions and the probability of adding a trustee fund is less sensitive to performance than the probability of adding a non-trustee fund. Interestingly, mirroring our results for deletions, we also find that addition probabilities are inversely related to performance among poorly performing trustee funds.

The trustee's tilt toward affiliated funds need not affect plan participants however. Although the investment opportunity set of the plan is determined by the menu choices selected by the employer and the trustee, participants can freely allocate their contributions within the opportunity set. If participants anticipate trustee biases or are simply sensitive to poor performance, they can undo favoritism in their own portfolios by, for instance, not allocating capital to poorly performing trustee funds. Therefore to investigate whether trustee favoritism has an impact on the overall allocation of plan assets, we examine the sensitivity of participant flows to the performance of trustee and non-trustee funds. We show that participants are generally not sensitive to poor performance and thus they do not undo the trustee bias. This in turn indicates that plan participants are affected by the trustee's behavior.

Finally, while our evidence on favoritism is consistent with adverse trustee incentives, trustees are also likely to have private information about their own proprietary funds. Therefore, it is possible that they show a strong preference for these funds in menu altering decisions not because they are biased toward them, but rather, due to positive information they possess about these funds. To investigate this possibility, we examine future fund performance. For instance, if – despite the lackluster past performance – the decision to keep poorly performing trustee funds on the menu is information driven, then they should perform better in the future. We find that this is not the case: trustee funds that rank poorly based on past performance but are not delisted from the menu do not perform well in the subsequent year. We estimate that on average they underperform by approximately 3.6% annually on a risk-adjusted basis. This figure is large in and of itself, but its economic significance is magnified in the retirement

context by compounding. Our results suggest that the trustee bias we document in this paper has important implications for the employees' income in retirement.

Our study belongs to a nascent literature on trustees in defined contribution plans. Davis and Kim (2007) and Cohen and Schmidt (2009) study conflicts of interest that exist in the trustee relationship. Both papers argue that to protect the valuable business relation that arises between the sponsoring company and the trustee family, fund families cater to the sponsors while compromising their own fiduciary responsibilities. In particular, Cohen and Schmidt (2009) find that trustee mutual fund families overinvest in the sponsoring company's stock. They also show that when other mutual funds sell the stock, trustee funds tend to trade in the opposite direction thereby creating liquidity for these shares and supporting the stock price of distressed firms. Davis and Kim (2007) document that the way trustee funds vote in shareholder meetings is influenced by the business ties they have as a trustee.

While these studies find that the sponsor can draw important benefits from appointing mutual fund trustees, mutual funds may also enjoy positive externalities from these arrangements in addition to capturing the large 401(k) asset pool. One such externality may be the opportunity to gain an information advantage about the sponsoring company through the trustee's access and connection to the company's management. Duan, Hotchkiss, and Jiao (2012) show that trustee families may indeed obtain valuable information about the sponsor, which provides them with profitable trading opportunities. We contribute to this literature by highlighting how adverse trustee incentives affect the fund lineup of the menu.

Our paper is also related to two additional areas of study. First, we contribute to the broader literature that focuses on the design and characteristics of defined contribution plans.³

³Papers on the design of employer-sponsored retirement accounts include Benartzi and Thaler (2001); Madrian and Shea (2001); Choi et al. (2002, 2004); Del Guercio and Tkac (2002); Duflo and Saez (2002); Agnew, Balduzzi, and Sunden (2003); Elton, Gruber, and Blake (2006, 2007); Brown, Liang, and Weisbenner (2007); Huberman and Jiang (2006); Rauh (2006); Goyal and Wahal (2008); Carroll et al. (2009); Tang et al. (2010); Balduzzi and Reuter (2012); Brown and Harlow (2012); Christoffersen and Simutin (2012); Mitchell

While papers in this literature commonly employ data that are either limited to plans offered by a single trustee or, alternatively, to a single year snapshot of the industry, our database contains a large cross-section of plans, trustees, and sponsors as well as an eleven-year time-series providing a rich laboratory for 401(k) research.

Second, our paper is related to the mutual fund literature on favoritism within fund families. Gaspar, Massa, and Matos (2006) show that mutual fund families strategically transfer performance across member funds to favor those funds that are more likely to increase overall family profits. They show that family organization generates distortions in delegated asset management. Reuter (2006) provides evidence that lead underwriters will use allocations of underpriced IPOs to reward those institutions with which they have strong business relationships. Kuhnlen (2009) investigates whether business networks mitigate agency conflicts by facilitating efficient information transfers or whether they are channels for inefficient favoritism. She finds that fund directors and advisory firms that manage the funds hire each other preferentially based on the intensity of their past interactions. Our paper provides evidence that mutual fund families favor their own affiliated funds when they act as trustees of 401(k) pension plans.⁴

The rest of the paper is structured as follows. Section 2 describes our data collection and provides summary statistics of our 401(k) plans as well as the mutual funds offered in the plans' menu. Sections 3–6 discuss our results. Section 7 concludes.

and Utkus (2012); Sialm and Starks (2012); and Sialm, Starks, and Zhang (2012).

⁴Additional papers studying favoritism within asset management companies include Nanda, Wang, and Zheng (2004); Ritter and Zhang (2007); Massa and Rehman (2008); Evans (2010); Bhattacharya, Lee, and Pool (2012); and Chaudhuri, Ivkovic, and Trzcinka (2012).

2 Data and Summary Statistics

This section describes the sample selection process and provides summary statistics for our sample of 401(k) menus.

2.1 Data collection

We collect the investment options offered in 401(k) plans from Form 11-K filed with the U.S. Securities and Exchange Commission (SEC). All plans offering the company stock as an investment option for plan participants are required to file this form with the SEC. The filing provides an overall description of the plan, identifies the trustee, all individual choices available to participants (the menu), and the accumulated value of assets invested in each of these vehicles at the end of the fiscal year. We manually collect these tables as disclosure is not standardized across plans and firms.

We start by webcrawling the SEC's website from 1998 to 2009 to identify all companies that report Form 11-K. We collect 26,624 links to 11-K filings but restrict this sample to companies covered by COMPUSTAT. We eliminate filings that have been submitted to the SEC in duplicate and consolidate amendments with the corresponding original filings. From these documents we collect all tables that describe the "Schedule of Assets" of the plan. In most cases, the table reports the complete set of investment options offered by the plan, including the employers' own stock, other common stocks, mutual funds, or commingled trusts, as well as the current value of investments in these options at the end of the fiscal year. Occasionally, the table describes only those investment options that capture more than 5% of the plan's assets or alternatively, only mutual fund investments. We supplement our Form 11-K information with plan level data from Form 5500, as described below. The resulting dataset has more than 302,000 observations, containing information at the firm-year-plan-menu level.

To obtain information on the characteristics of the mutual funds included in DC plans, we match all funds listed on the menus to the CRSP Survivorship Bias-Free U.S. Mutual Fund database. To aid our matching task, we proceed in several steps. We start by filtering our menu options for non-mutual fund assets. These include, for instance, common stocks, bonds, or guaranteed investment contracts. In approximately 20% of the cases, the SEC Form 11-K contains information on the number of shares of each asset held by the plan in addition to the market value of the position. This allows us to calculate the net asset value (NAV) of the position on the report date. When the NAV information is available, we match the menu choice to the CRSP mutual fund files by NAV and date. For the rest of the sample, we hand match the 11-K funds to the mutual fund database by name.

Since most plans do not identify the exact share class of the fund offered on the menu, we establish the link between our 401(k) sample and the CRSP Survivorship Bias-free Mutual Fund database at the fund-level, that is, we combine information on all available share classes of each fund in CRSP into fund-level variables. Accordingly, fund age is calculated as the age of the oldest share class, fund size is the sum of the total net assets of all share classes, and fund returns and expense ratios are calculated as the total net asset value weighted average returns and expense ratios of the share classes, respectively. We also classify each mutual fund in our sample as “balanced,” “bond,” “domestic equity,” “international equity,” or “other.” We create separate dummy variables for money market funds, target date funds, and index funds. We manually group funds into target date and index fund categories based on fund name. Around 62% of the funds in the average plan in our sample are equity funds and 20% are bond funds. There is a steady increase in the number of target date funds over our sample period, especially after the passage of the Pension Protection Act (PPA) of 2006, also documented by Mitchell and Utkus (2012).⁵

⁵Following the PPA (2006), the Department of Labor added a new fiduciary protection to ERISA for default

Finally, we perform two additional data steps to complete our sample. First, we assign unique plan IDs to create time-series at the plan level. Form 11-K does not always disclose the plan number. Companies occasionally sponsor multiple plans for different subsidiaries, salaried and hourly employees, or unionized and non-unionized workers. In order to track the same plan over time, we collect the plan Employer Identification Number (EIN) and Plan Number (PN) by searching Form 5500 by plan name and assets. Once established, the link with Form 5500 allows us to collect additional information on total participants, active participants, employer and employee contributions, total assets, and whether the plan is collectively bargained or not.

We manually collect the trustee name (and any trustee change occurring during the year) from the plan description available in Form 11-K. We supplement and cross check this information with the name of the trustee disclosed in Form 5500.

2.2 Sample description

Table 1 describes the composition of our sample by year. Our data covers 2,645 distinct plans sponsored by 1,853 firms from 1998 to 2009 (companies can sponsor multiple plans). Overall, the final dataset has 13,585 plan-year observations. The number of plans is smaller during the early part of the sample as disclosures on investments were generally less standardized. Similarly, our data for 2009 are potentially incomplete as they do not include late filers or filers who have a late fiscal year end. More than 56% of the menu options are mutual fund choices.⁶

investments. Section 404(c)(5) stipulates that if participants wave the right to direct their investments, the fiduciaries (sponsors and trustees) are protected from suits about the default investment option's market performance if those participants are invested in a qualified default investment alternative (QDIA). Eligible QDIAs include target-date funds, traditional balanced funds, and managed account advice services.

⁶For robustness, we also run our analyses below by excluding the first and the last years of our sample. This has no significant effects on our results.

The table also provides information about plan trusteeship and architecture. About 77% of plans in our sample are trusteeed by a mutual fund family. The remaining plans are trusteeed by commercial banks, consulting companies, individuals, or by the sponsoring company itself. We do not differentiate between these later categories and collapse them into one group which we refer to as “Other Trustees” or “Non-Mutual Fund Trustee.” The ratio between the number of plans trusteeed by management companies and by other entities has been slightly increasing over time but experienced a small decrease in the later years as a response to the increased competition in this market.⁷

The table shows an increase in the number of mutual fund investment options offered in the average plan over time. To summarize the growing importance of including funds from a number of different mutual fund families on the menu, which we refer to as *open architecture*, we report three metrics. *Trustee share* represents the proportion of total plan assets invested in mutual funds that are offered by the trustee family. Since this is zero for plans with non-mutual fund trustees, trustee share in this table provides the overall proportion of retirement assets invested with affiliated funds. We also report the average number of management companies that offer at least one investment option on the menu and the Herfindahl index of the menu calculated based on the dollar share of each of these management companies. These measures indicate a decline in the share of the assets managed by trustee families and an increase in the popularity of offering mutual funds from several different families.

In our sample, the average size of a 401(k) plan is approximately \$280 million (with a median of \$60 million), suggesting that our dataset covers several very large, well-established plans. The average plan size generally increases over our sample period, peaking at \$364 million in 2007. On aggregate, our plans cover \$376 billion in retirement assets in 2009

⁷See for instance, the Bloomberg article on March 23, 2011: <http://www.bloomberg.com/news/print/2011-03-23/banks-angle-for-bigger-share-of-4-trillion-retirement-market.html>

(\$485 billion in 2007) and over 9 million total participants per year. The typical account size is \$41,365 and employees contribute \$5,200 per year. The percentage of assets invested in employer securities varies across plans and years, with a mean of 17% of assets and a median of 10%. About 13% of our plans are collectively negotiated (i.e. unionized). Our sample is representative of the universe of plans sponsored by public companies filing Form 5500 with the Department of Labor in terms of plan size, number of participants, and industry composition.⁸

Table 2 describes the characteristics of mutual funds that have been added, kept, or deleted from the menu, by their trustee affiliation. We also report the difference between trustee and non-trustee characteristics along with the results of paired *t*-tests that evaluate the statistical significance of the difference. The corresponding standard errors are two-way clustered at the plan and fund levels.

Our sample contains 19,003 fund deletions, 139,182 fund-year observations that stay in the sample for at least two consecutive years, and 28,193 fund additions. Overall, funds that are deleted from the plans in our sample have the lowest average performance across the three groups, as measured by their percentile performance rank among funds of the same style in the CRSP mutual fund universe using past one-, three-, and five-year returns. Added funds are younger and come with better performance records than those that are kept or deleted from the menus. Trustee funds that are kept or added generally have lower performance over the past three and five years.

The table also shows that fees charged by trustee funds are, on average, significantly lower than those of non-trustee funds. Since in most cases we are not able to identify the exact

⁸Our sample covers 30-35% of the 401(k) assets of plans sponsored by publicly listed companies that report Form 5500. Filing Form 5500 with the Department of Labor (DOL) and Internal Revenue Service (IRS) is mandatory for all employee benefit plan that qualify under the Employee Retirement Income Security Act of 1974 (ERISA).

share class of the fund offered in the plan, we calculate fees as the weighted average expense ratio of all share classes reported in CRSP with the total net assets of each share class as weights. The difference between the expense ratios of trustee and non-trustee funds may be driven by the difference in fund styles across the two groups. To investigate this possibility, we also report average style-adjusted fees. We calculate the style-adjusted average expense ratio by subtracting from each fund's expense ratio the value weighted average expense ratio of all funds in its investment category as determined by its Lipper Objective Code. We find that trustee funds are cheaper on a style-adjusted basis as well.⁹

These results point to a potential benefit of employing mutual fund trustees, as they typically offer their own low-fee funds. Nevertheless, and probably the least explored dimension of this relationship in the literature, is the potential cost engendered by trustee favoritism. This paper investigates this hypothesis.

3 Fund Deletions

Investment allocations in 401(k) accounts are driven by both plan providers and plan participants. In a first step, the trustee, along with the sponsor and other plan fiduciaries, determines the investment options for the plan. In a second step, plan participants allocate their retirement savings and contributions to the various investment options. It is also the responsibility of the provider to insure that the plan continuously offers a suitable set of choices. Therefore, trustees dynamically adjust 401(k) menus by deleting some investment options and adding others. In this section, we study whether plan trustees favor their own mutual funds relative to funds affiliated with other mutual fund companies during these menu altering decisions.

⁹In unreported analyses, we find that the expense ratio of trustee funds is significantly lower even after controlling for fund size, age, turnover, as well as various plan characteristics.

3.1 Univariate Relationship

We first provide univariate analyses to investigate whether the propensity to delete a fund from the menu depends on whether the fund is affiliated with the trustee. A mutual fund can be an investment option in a 401(k) plan where its management company is the plan's trustee ("trustee fund"), as well as in a plan offered by an unaffiliated trustee ("non-trustee fund"). Therefore, in each year, for each fund, we count the number of menus on which the fund is a trustee fund and the number of menus on which it is a non-trustee fund, respectively. We then count the number of affiliated and the number of unaffiliated menus from which the fund is delisted during the year. This allows us to determine the deletion frequencies for each fund for each year by affiliation.

To make the comparison between the deletion frequencies of trustee and non-trustee funds more meaningful, we also group funds into deciles based on past performance. In particular, we compute the style-adjusted returns of all mutual funds in the CRSP mutual fund database over the prior one, three, and five years. Funds are then sorted into decile portfolios based on their style-adjusted performance.

Figure 1 reports the mean annual deletion frequencies by trustee affiliation for each performance decile using the prior 36 months to evaluate performance. We construct the figure by first computing the average deletion rates for each fund in each year in affiliated and unaffiliated 401(k) plans, as described above. In a second step, we average the deletion rates within the performance deciles by year. Finally, we average the decile deletion rates across time. Panel A shows the results based on all funds in our sample. In Panel B, we focus only on those funds that contemporaneously appear on multiple 401(k) menus, at least once as a trustee fund and at least once as a non-trustee fund. By comparing the deletion probabilities of the same fund across plans managed by different trustees, our results are not contaminated

by different fund characteristics or performance records.

The figure shows that a fund is significantly less likely to be deleted from a plan in which it is a trustee fund than from those in which it is a non-trustee fund regardless of past performance. For example, Panel A indicates that the average trustee fund has a deletion rate of 11.6% across all performance deciles, whereas a non-trustee fund has an average deletion rate of 21.4%. Furthermore, we find that the differences in deletion rates widen significantly if we focus on poorly performing funds. For example, funds in the lowest performance ranking decile in Panel A have a probability of deletion of 29.6% for non-trustee funds and a probability of deletion of only 11.9% for trustee funds. Indeed we observe that the deletion rate of trustee funds in the lowest performance decile is actually lower than the deletion rates of trustee funds in deciles two through four. This is surprising provided that Carhart (1997) documents performance persistence among poorly performing funds.

Panel B shows similar results for the subsample of funds that are simultaneously offered as both trustee and non-trustee funds. In this analysis the funds in each decile are identical across the affiliated and unaffiliated groups. Thus, our results are not driven by differences in fund characteristics.

Table 3 summarizes the corresponding deletion frequencies by performance deciles based on one, three, and five year evaluation horizons. The table reports the deletion frequencies for trustee funds and non-trustee funds, as well as the difference between them. Analogously to Figure 1, average frequencies in Panel A are based on all funds in our sample, while Panel B calculates deletion probabilities using only those funds in our sample that simultaneously appear as trustee and non-trustee funds. Overall, we find that a fund is significantly more likely to be deleted from those menus on which it is not a trustee fund regardless of past performance under all three performance ranking horizons. Moreover, the difference between deletion probabilities is largest for poorly performing funds.

3.2 Multivariate Relation

Restricting our attention to those funds that simultaneously appear on several different menus offered by different trustees allows us to hold fund characteristics constant in our univariate analyses. To investigate whether the results in Section 3.1 are robust to controlling for various menu characteristics and to examine the performance sensitivity of affiliated and unaffiliated fund deletions, we estimate the following model:

$$\begin{aligned}
 DEL_{p,f,t} &= \beta_0 + \beta_1 \times TF_{p,f,t} + \beta_2 \times LowRank_{p,f,t} + \beta_3 \times HighRank_{p,f,t} \\
 &+ \beta_4 \times TF_{p,f,t} \times LowRank_{p,f,t} + \beta_5 \times TF_{p,f,t} \times HighRank_{p,f,t} \\
 &+ Z'_{p,f,t} \gamma + \epsilon_{p,f,t},
 \end{aligned} \tag{1}$$

where $DEL_{p,f,t}$ is an indicator variable that takes the value of one if mutual fund f has been deleted from plan p at time t and zero otherwise, and $TF_{p,f,t}$ is an indicator variable for whether the trustee of pension plan p is affiliated with the management company of mutual fund f . $LowRank$ and $HighRank$ are defined as $LowRank_{p,f,t} = \min(Rank_{p,f,t}, 0.5)$ and $HighRank_{p,f,t} = \min(Rank_{p,f,t} - LowRank_{p,f,t}, 0.5)$, where $Rank_{p,f,t}$ is the performance rank of mutual fund f over the previous one, three, or five years. Performance ranks are formed based on the style-adjusted returns of all mutual funds in the CRSP mutual fund database over the prior one, three, and five years, as described in the previous section.

The other control variables Z include the natural logarithm of the option size (plan assets invested in the fund), the number of options, the expense ratio of the fund, the turnover of the fund's holdings, the natural logarithm of the fund's size, fund age, the standard deviation of the fund's return, and unreported indicator variables for specific fund types (e.g., domestic equity, international equity, balanced, bond, target date, index, and money market funds) and time (year) fixed effects.¹⁰

¹⁰While fund-year fixed effects would allow for identification based on the same fund appearing on multiple

In our baseline model described in equation (1), we use two performance segments, evaluating the trustee's response to performance separately for below median and above median ranks. For robustness however, we also estimate our model using quintile based performance segments following Sirri and Tufano (1998) as well as a one-segment (linear) model. Favoritism toward affiliated funds implies that, all else equal, trustee funds are less likely to be delisted (i.e., $\beta_1 < 0$) and that trustee deletions are less sensitive to poor prior performance (i.e., $\beta_4 > 0$).

Table 4 reports the coefficient estimates. We estimate equation (1) using a linear probability model, which allows for a straightforward interpretation of the piecewise linear terms and the corresponding interactions. The standard errors in the table are two-way clustered at the plan and fund levels.¹¹ Consistent with Figure 1 we find that trustee funds are significantly less likely to be deleted. A trustee fund has a lower probability of being delisted that ranges between 9.9% and 14%, depending on the specification. Consistent with performance chasing, we find that the probability of deletions decreases significantly with fund performance. For example, a ten percentage point increase in the fund's performance rank decreases the probability of deletions by between 1.8% and 3.2% for below median funds. Finally, we also find that deletions of trustee funds are less sensitive to poor performance than non-trustee funds as indicated by the highly significant positive β_4 coefficient. For trustee funds, the sensitivity of deletions to inferior fund performance is less than half of that of non-trustee funds.

The additional control variables indicate that funds with large plan investments are less

menus, as in our univariate setup, a fund's performance only varies over time. Therefore, to estimate the performance sensitivity of fund deletions, we do not use fund-year fixed effects.

¹¹In Table A1 of the Appendix, we report the corresponding estimation results using a probit specification. The table displays the estimated marginal effects. For the interaction terms, these are calculated using the INTEFF command based on Ai and Norton (2003) and correspond to the average estimated marginal effect. Figure A1 of the Appendix provides a more complete picture and displays the individual marginal effect estimates of the interaction terms for each observation of our sample along with the corresponding z -statistics. The findings in the table and the corresponding inteff graphs are qualitatively identical to those in Table 4. In the rest of the paper, we only report our estimates using a linear probability model.

likely to be deleted and that plans with more investment options are less likely to delete a specific fund. Plan providers are also more likely to delete funds with high expense ratios, funds with high turnover, and smaller funds. Overall, our baseline results indicate that trustee funds are significantly less likely to be deleted from 401(k) plans than non-trustee funds and that this bias is particularly pronounced for poorly performing funds.

3.3 Robustness Tests

In this section, we report additional robustness tests for the base-case results summarized in Section 3.2.

3.3.1 Alternative Performance Ranking

In our baseline specification we rank mutual funds according to the style-adjusted returns of all mutual funds in the CRSP mutual fund database. We refer to this global ranking as “Overall Ranking.” For robustness we now compute two alternative ranking methods, where the percentile performance rank of a fund is either measured relative to the other investment options in a specific 401(k) plan (“Plan Ranking”) or relative to the other funds offered by the fund’s family (“Family Ranking”). The overall ranking method captures the performance of a fund relative to the universe of available mutual funds in the U.S., which could be viewed as the most comprehensive metric. When a fund underperforms compared to the other investment choices included in the plan or the other options in the fund family, the trustee may be pressured to remove the fund from the menu as underperformance in this setting is perhaps more transparent.

Table 5 summarizes the coefficient estimates from equation (1) when $Rank_{p,f,t}$ is defined using the alternative ranking methodologies. The results are qualitatively and quantitatively similar to the base-case results reported in Table 4. Thus, our findings are not affected by

whether we benchmark mutual funds relative to the universe of mutual funds or relative to other funds included in the same 401(k) plan or other funds offered by the same fund family.

3.3.2 Linear Performance

In columns 1-3 of Table 6, we reestimate our baseline regression using a linear performance model (single performance segment) for robustness. For brevity, the table only reports the results using the three year horizon. Column 1 is based on performance ranking relative to funds in the same objective codes (i.e., overall ranking) and columns 2 and 3 report the corresponding results using the fund's performance rank relative to the 401(k) plan or the fund family. Consistent with the base-case specification from Table 4, we find that trustee funds are significantly less likely to be deleted with the difference in probabilities ranging from 5.1% to 9.5%. As in our baseline results, the sensitivity of trustee fund deletions to performance is significantly smaller for trustee funds.

3.3.3 Sensitivity to Extreme Performance

To analyze in more depth the sensitivity of deletions to extreme performance, we estimate a specification using three piecewise linear segments instead of the two segments from equation (1). The performance segments are 1) the lowest performance quintile, 2) the highest performance quintile, and 3) the three middle performance quintiles, which are pulled together to represent a single performance segment. Following Sirri and Tufano (1998), the performance in the lowest quintile is given by $LowQRank_{p,f,t} = \min(Rank_{p,f,t}, 0.2)$, the performance in the three middle quintiles is given by $MidQRank_{p,f,t} = \min(Rank_{p,f,t} - LowQRank_{p,f,t}, 0.6)$, and the performance in the highest quintile is given by $HighQRank_{p,f,t} = (Rank_{p,f,t} - LowQRank_{p,f,t} - MidQRank_{p,f,t})$.

Columns 4-6 of Table 6 report the estimates from the three piecewise linear segments

using our alternative ranking methods, based on the three year horizon. Consistent with the base-case specification from Table 4, we find that deletions are less sensitive to poor and intermediate performance for trustee funds. Interestingly, in our overall ranking model, we find that the deletions of non-trustee funds that rank in the highest performance quintile relative to other funds in the same objective codes actually increase with the performance rank.

3.3.4 Subsample Analysis

Table 7 shows the results of our linear probability model specified in equation (1) for various subsamples. For brevity, we only report the results using the three year horizon with the overall performance ranking. In the first four columns, we confirm that our results are consistent across different trustees. In the first column, we exclude the three largest trustees each year. These are the only trustees in our sample that have over 10% of all retirement assets. In column 2, we report our estimates for the three largest trustees only. Overall, we find qualitatively similar results across the two subsamples. Whereas affiliated trustee funds are 13.9% less likely to be deleted from the menu for non-top mutual fund trustees, we find that trustee funds are 12.7% less likely to be deleted from the menu for top-three mutual fund trustees. In addition, we find that trustee fund deletions are less sensitive to poor fund performance for both large and small mutual fund trustees. In column 3, we include trustee fixed effects since deletion probabilities might depend on the identity of the trustee. Our favoritism results remain after controlling for trustee fixed effects. In column 4, we reestimate our results using information only on those plans that are trustee by a mutual fund family.

Columns 5 and 6 restrict the sample of mutual funds considered. In column 5 we exclude all target date funds, since these funds are often used as default investment options. In column 6, we restrict our sample to equity funds. The results in these specifications are very consistent

with the results in our base-case specification.

The Pension Protection Act of 2006 (PPA) introduced comprehensive new legislation for U.S. pension plans to protect retirement plan participants. Although the majority of the reforms concerned defined benefit plans, it also affected defined contribution plans by allowing companies to offer objective investment advice to participants and by requiring plans to provide specific benefit statements to participants.¹² Furthermore, several class action lawsuits were filed in the mid 2000s against large employers for breaches of fiduciary obligations with respect to their 401(k) accounts.¹³ To investigate whether these lawsuits and regulatory reforms affect our results, we divide our sample into two subperiods (1998-2006 and 2007-2009). Columns 7 and 8 of Table 7 indicate that our key results do not differ between the two subperiods. We find that trustee funds exhibit a lower propensity to be deleted from 401(k) menus and that deletions for trustee funds are less sensitive to prior fund performance for both subperiods.

4 Fund Additions

The previous section provides evidence that trustees are substantially less likely to delete their own funds from the menus, and even more so when these funds are poorly performing. In this section, we document that similar biases exist for fund additions as well. We first provide univariate analyses to investigate whether the propensity to add a fund to a menu depends on the fund's affiliation with the trustee. We then examine the characteristics of newly listed trustee and non-trustee funds in a multivariate framework.

¹²The detailed regulations from the 2006 Pension Protection Act can be obtained from <http://www.dol.gov/ebsa/pensionreform.html>.

¹³See Ruiz-Zaiko and Williams (2007) for additional information on the lawsuits.

4.1 Univariate Relationship

We begin our analyses by calculating addition frequencies for affiliated and unaffiliated investment options, respectively. In Section 3.1, we compute mean deletion frequencies by averaging across the deletion propensities of all funds that could be deleted from a 401(k) plan in a given year. Our sample of 401(k) investment options represents the set of these funds. Analogously, to calculate mean addition frequencies we first determine the addition propensities of each fund that could be added to a 401(k) menu in a given year. For additions, the set of funds that could be added is represented by the CRSP mutual fund universe.

To investigate how a fund's propensity to be added to a menu depends on its affiliation with the trustee, for each fund in CRSP we determine its addition frequency as an affiliated and unaffiliated menu choice, respectively. The affiliated (unaffiliated) addition frequency of a fund is defined as the number of affiliated (unaffiliated) plans to which the fund is added as a new investment option during the year divided by the number of affiliated (unaffiliated) plans in which the fund is not already offered as an option at the end of the previous year (i.e., the total number of affiliated (unaffiliated) menus to which it could be added).

Consistent with Figure 1 and Table 3 for fund deletions, each year we also sort funds in the CRSP universe into deciles according to their style-adjusted performance over the prior one, three, and five years. Figure 2 and Table 8 report the average affiliated and unaffiliated addition frequencies by past performance decile. In both cases, Panel A summarizes the results using all existing mutual funds, whereas the average frequencies in Panel B are based on funds from only those families that act as a trustee for at least one of our 401(k) plans during the year. Thus, Panel B excludes funds that could not be added as trustee funds during the year. This restriction allows us to examine the addition frequency of the same fund to an affiliated or unaffiliated menu, respectively, consistent with our analyses for deletions in Section 3.1

above.

Figure 2 shows a substantial difference between the average addition frequencies for trustee and non-trustee funds using past performance rankings based on the previous 36 months. In the overall sample, the average addition frequency equals 1.08% for trustee funds and just 0.02% for non-trustee funds. Thus, trustee funds are more than fifty times more likely to be added to a plan than non-trustee funds, indicating a substantial trustee bias for fund additions.¹⁴ This result extends to using performance rankings based on the previous one and five years, as summarized in Table 8.

We also find that addition frequencies increase disproportionately more with fund performance for non-trustee funds than for trustee-funds, indicating that non-trustee additions are much more sensitive to performance. An improvement in performance from the lowest to the highest decile increases the addition probability for non-trustee funds approximately eight-fold from 0.005% to 0.039%. At the same time, an equivalent improvement in performance for trustee funds results in only a 2.5 times larger addition probability (from 0.69% to 1.68%).

Finally, the results indicate that trustee funds in the bottom performance decile are more likely to be added to 401(k) plans than trustee funds in the second performance decile. For example, trustee funds in the lowest performance decile over the prior three years have an addition frequency of 0.69%, whereas trustee funds in the second performance decile have an addition probability of just 0.47%. In contrast, we do not find such non-monotonicities for non-trustee funds. These results are consistent with the hypothesis that plan trustees support their affiliated poorly-performing funds by adding them to 401(k) plans.

Panel B of Figure 2 and Table 8 replicates these results for the subsample of funds offered by families that serve as trustees of 401(k) plans. The affiliated addition frequencies are

¹⁴The difference in addition frequencies is similarly stark when we limit our analysis to only those investment styles in the CRSP universe that appear on 401(k) menus in our sample.

exactly identical between the two panels. However, the addition frequencies of non-trustee funds are slightly higher for the subsample in Panel B since funds offered by families that serve as trustees are more likely to be selected by other unaffiliated trustees for their 401(k) plans.¹⁵

4.2 Sample of Additions

Next, we investigate the characteristics of affiliated and unaffiliated funds based on our sample of newly added funds. Table 2 from Section 2 provides univariate evidence that newly listed trustee funds exhibit lower past performance than non-trustee funds in the same category. We confirm this finding in Figure 3. The figure describes the distribution of trustee and non-trustee fund additions separately, by performance deciles. Fund performance is measured by the percentile performance rank of each fund in its style category in the past one, three, and five years, respectively, using the CRSP universe of mutual funds. The graph shows results based on fund performance in the past three years, but using one or five year ranks produces qualitatively similar patterns. The results reveal that the proportion of non-trustee funds with strong past performance is larger compared to that of trustee funds, while trustee funds are more likely to come to the menu with a mediocre performance record.

To further explore the difference in past performance across newly added trustee and non-trustee funds, we estimate the following linear probability model for trustee additions:

$$TF^{ADD}_{p,f,t} = \beta_0 + \beta_1 \times Rank_{p,f,t} + Z'_{p,f,t}\gamma + \epsilon_{p,f,t}, \quad (2)$$

where the dependent variable takes the value of one if fund f added to plan p at time t is a trustee fund, and zero otherwise. Since the sample used in this analysis includes only

¹⁵In an unreported robustness test we estimate the addition proportions for trustee and non-trustee funds using only mutual funds that are included in at least one 401(k) plan. Although the addition proportions are around three times higher using this fund subsample, the results are qualitatively unaffected relative to the base-case specification.

fund additions, it reflects the choice between selecting a trustee fund over a non-trustee fund. $Rank_{p,f,t}$ is the percentile performance rank of mutual fund f over the previous one, three, or five years based on overall rankings and it enters the analysis as a linear term. Our additional controls include various fund characteristics and plan level variables, such as the number of menu options and plan size.

The results are reported in Table 9 with standard errors two-way clustered at the plan and fund levels. Consistent with trustee favoritism, trustee fund additions are associated with worse past performance even after controlling for other fund characteristics. This is represented by our $Rank_{p,f,t}$ coefficient estimates, which are significantly negative at the one percent level for each of our performance measures.¹⁶

5 Participant Flows

While the investment opportunity set of the plan is determined by the menu choices selected by the employer and the trustee, participants can freely allocate their contributions within the opportunity set. If participants anticipate trustee biases or are simply sensitive to poor performance, they can undo favoritism in their own portfolios by, for instance, not allocating capital to poorly performing trustee funds that are not removed from the menu. In this section, we investigate whether trustee favoritism has an impact on the overall allocation of plan assets by examining the sensitivity of participant flows to the performance of trustee and non-trustee funds.

Our primary definition of the growth rate of new money of fund f held in 401(k) plan p at time t is based on the following definition of fund flows:

$$NMG1_{p,f,t} = \frac{V_{p,f,t} - V_{p,f,t-1}(1 + R_{f,t})}{V_{p,f,t-1}(1 + R_{f,t})}. \quad (3)$$

¹⁶These results are reported using trustee fixed effects but are qualitatively equivalent when the trustee fixed effects are not included.

The numerator captures the dollar change in the value of participants' investments ($V_{p,f,t}$) in fund f in plan p in year t after adjusting for the price appreciation of plan assets $R_{f,t}$ (i.e., fund return) during the year. The denominator is defined as the projected value of the lagged plan position in the fund without any new flow of money. If an investment option is deleted from a plan menu, then $NMG1$ equals exactly -100%. This definition of new money growth allows us to decompose fund flows to existing menu options¹⁷ into a component that is driven by the plan sponsor and the trustee (i.e., fund deletions, which are extensive margin flows of -100%) and a component that is driven primarily by plan participants (i.e., all other changes in the value invested in the menu option, which are intensive margin flows above -100%).¹⁸ The analysis of extensive margin flows representing the decisions of plan sponsors and trustees is summarized in Section 3.2. To remove outliers, we winsorize $NMG1$ at the 95% level.

Figure 4 depicts the histograms of the percentage flows into various plan options for trustee and non-trustee funds in the lowest performance quintile over the previous three years based on various ranking methodologies. Consistent with Figure 1, which focuses on a special subset of the funds in our sample, we find that non-trustee options are significantly more likely to be deleted than trustee options overall. Figure 4 also shows that deletions contribute significantly to the total flows of new money, consistent with Sialm, Starks, and Zhang (2012).

Since equation (3) is not defined for fund additions, we also adopt two alternative measures for the growth rate of new money of fund f held in 401(k) plan p at time t using two alternative denominators for equation (3). The denominator of our first alternative measure ($NMG2$) is

¹⁷ $NMG1$ is not defined for newly listed funds as for these, the value of the lagged plan position is 0.

¹⁸Whereas the extensive margin $NMG1$ rates are fund deletions, which are fully driven by choices of the plan sponsors and trustees, the intensive margin $NMG1$ rates are not only affected by plan participants but also by plan sponsors and trustees (e.g., additions of new funds that compete with old funds, preventing contributions to certain funds but grandfathering balances held by participants in these funds, selecting some funds as default investment options).

$V_{p,f,t} + V_{p,f,t-1}(1 + R_{f,t})$:

$$NMG2_{p,f,t} = \frac{V_{p,f,t} - V_{p,f,t-1}(1 + R_{f,t})}{V_{p,f,t} + V_{p,f,t-1}(1 + R_{f,t})}. \quad (4)$$

Under this definition, new money growth takes a value in the interval $[-1,1]$. In particular, it equals -100% for a fund that is eliminated as an investment option, as before, and +100% for a fund that is newly added to the pension plan. More gradual inflows and outflows into the fund are represented by intermediate values. Extensive margin new money growth that equals -100% or +100% corresponds to menu changes by sponsors and trustees. Intermediate values correspond to the changes plan participants make to their asset allocations.

Finally, the denominator of our second alternative new money growth measure ($NMG3$) is the plan value from the previous year:

$$NMG3_{p,f,t} = \frac{V_{p,f,t} - V_{p,f,t-1}(1 + R_{f,t})}{\sum_f V_{p,f,t-1}(1 + R_{f,t})}. \quad (5)$$

To remove outliers, we winsorize $NMG3$ at the 95% level. While the values $NMG3$ takes under this third definition are not restricted to a specific range, we separate participant actions from those of the sponsors and the trustees by removing the set of additions and deletions from that of the rest of our sample.

To investigate the sensitivity of fund flows to prior performance, we estimate the following OLS regression using the three alternative definitions of NMG :

$$\begin{aligned} NMG_{p,f,t} &= \beta_0 + \beta_1 \times TF_{p,f,t} + \beta_2 \times LowRank_{p,f,t} + \beta_3 \times HighRank_{p,f,t} \\ &+ \beta_4 \times TF_{p,f,t} \times LowRank_{p,f,t} + \beta_5 \times TF_{p,f,t} \times HighRank_{p,f,t} \\ &+ Z'_{p,f,t} \gamma + \epsilon_{p,f,t}. \end{aligned} \quad (6)$$

Equation (6) is analogous to our baseline equation (1) with two exceptions. First, our new dependent variable in equation (6) is NMG , a continuous variable under all three definitions

(which replaces fund deletions, i.e., extensive margin *NMG* of -100%). Second, if participants use the same allocation rule every year, growth occurs mechanically due to the additional money contributed to the retirement accounts over time. To capture this mechanical characteristic of intensive margin flows, we add plan growth as an additional control variable.¹⁹

Table 10 reports the corresponding coefficient estimates using the overall performance ranking based on the past 36 months. The first three columns report the coefficient estimates using the full sample of *NMG* including both extensive and intensive margin flows. The last three columns report the coefficient estimates using only the intensive margin *NMG* (including flows strictly larger than -100% for the first definition and including flows strictly between -100% and +100% for the second, for example).

The main results in columns 1–3 using total fund flows are broadly consistent with the deletion results from Table 4. Trustee funds attract more new money than non-trustee funds.²⁰ We find that flows into various plan options increase with prior fund performance consistent with Chevalier and Ellison (1997), Sirri and Tufano (1998), and Huang, Wei, and Yan (2007). The interaction effects indicate that overall flows are significantly less sensitive to poor performance for trustee funds. For example, a ten percentage point increase in the performance rank over the previous three years for below-median funds increases flows over the next year by 6% for non-trustee funds and by only 1% for trustee funds. The additional control variables indicate that the growth rates tend to be larger for smaller investment options, for funds with lower expense ratios, for larger funds, and for younger funds.

To investigate the importance of participant flows, we restrict our attention to the intensive

¹⁹We calculate plan growth, using information in Form 5500 on total contributions, total expenses and total assets.

²⁰Our framework is significantly different from the flow benefit analysis in Cohen and Schmidt (2009). While they show that funds offered by the trustees in 401(k) menus have generally higher inflows and lower outflows (both retirement and retail flows), our paper investigates the different treatment funds are subject to when they belong to the trustee and when they do not.

margin money flows in the last three columns of Table 10. We find that participant flows are at best only marginally higher for trustee funds. Thus, the higher overall flows to trustee funds are primarily driven by the decisions of plan trustees and sponsors. The coefficients on the two performance ranking segments indicate that participants chase prior fund performance. It is interesting that most of the inflows into above-median performers are due to plan participants, whereas most of the outflows out of below-median performers are due to plan trustees. The interaction effects between trustee funds and performance ranks indicate that plan participants do not differentiate much between trustee and non-trustee funds.

Overall, we find that plan participants do not offset the biased decisions of plan sponsors and trustees. Our results indicate that decisions of plan trustees have a substantial impact on flows to mutual funds. Mutual fund trustees can benefit by obtaining higher money flows into their affiliated funds and by avoiding large outflows for their poorly performing funds.

6 Future Performance

Our previous results indicate that 401(k) plan sponsors are less likely to delete trustee funds from their menus and that deletions of trustee funds are less sensitive to prior fund performance. We also document a similar behavior for fund additions. Finally, we show that participants do not direct flows away from the biased options offered by the trustee.

Still, favoritism toward affiliated funds may not hurt plan participants if the underperforming trustee funds exhibit superior subsequent performance. Indeed trustees may keep poor performers not because they are biased toward them, but rather, due to positive information they possess about the future returns of these funds.

To investigate this hypothesis, in this section we examine the performance of trustee and non-trustee funds that are kept, deleted, or added in the plans. At the end of each calendar year, we form equal-weighted portfolios of trustee and non-trustee funds separately based on

whether the funds were kept, deleted, or added to the 401(k) menu (“No Changes,” “Deletions,” and “Additions”) during the calendar year. We restrict our sample to domestic equity funds in these analyses. This creates six portfolios. We then further subdivide these six groups based on past performance. In particular, “All Funds,” refers to all funds in the original six portfolios and “Lowest Quintile,” (“Lowest Decile”) refers to a subportfolio in each group that contains only those funds that also rank in the lowest performance quintile (decile) based on past performance. We use the overall performance rankings during the prior three years as our baseline specification. For example, “Trustee Funds/Deletions/All Funds” refers to the equally-weighted portfolio of all trustee funds that are deleted from a menu during the year, while “Non-trustee Funds/Deletions/Lowest Decile” represents the portfolio of poorly performing non-trustee funds that are deleted from a menu. We rebalance our portfolios at the end of each year and calculate the portfolios’ return for each of the next 12 months keeping the portfolios’ composition fixed.

The abnormal return $\alpha_{f,t}$ of fund portfolio f at time t is computed using the Fama-French-Carhart four-factor model (FFM) over our complete sample period using monthly fund return data from the CRSP Mutual Fund database:

$$\begin{aligned}
 R_{f,t} - R_{TB,t} = & \alpha_{f,t} + \beta_{f,t}^M (R_{M,t} - R_{TB,t}) + \beta_{f,t}^{SMB} (R_{S,t} - R_{B,t}) \\
 & + \beta_{f,t}^{HML} (R_{H,t} - R_{L,t}) + \beta_{f,t}^{UMD} (R_{U,t} - R_{D,t}) + \epsilon_{f,t}. \quad (7)
 \end{aligned}$$

The return of fund portfolio f during time period t is denoted by $R_{f,t}$. The index M corresponds to the market portfolio and the index TB to the risk-free Treasury bill rate. Portfolios of small and large stocks are denoted by S and B , respectively; portfolios of stocks with high and low ratios between their book values and their market values are denoted by H and L , respectively; and portfolios of stocks with relatively high and low returns during the previous year are denoted by U and D , respectively. We obtain monthly factor returns and the

risk-free rate from Kenneth French's website. The Carhart (1997) model nests the CAPM model (which includes only the market factor) and the Fama and French (1993) model (which includes the size and the book-to-market factors in addition to the market factor).

Table 11 reports the abnormal returns of the various mutual fund portfolios. Panels A, B, and C report the Carhart alphas, the Fama-French alphas, and the CAPM alphas, respectively. The average Carhart alpha for trustee funds kept for at least two consecutive periods in the 401(k) plan is -0.04% per month and is not significantly different from zero. Similarly, the corresponding alpha for non-trustee funds is -0.06% per month.

More importantly, we find that trustee funds that were kept in the 401(k) plans by their sponsors despite their poor performance exhibit significantly negative Carhart and Fama-French alphas. For example, trustee funds ranked in the lowest performance quintile (decile) over the prior three years exhibit a Carhart alpha of -0.24% (-0.30%) per month. The results using the Carhart and the Fama-French alphas are both statistically and economically significant. On the other hand, the results are less pronounced using CAPM alphas. This represents an underperformance of between 2.9% and 3.6% per year, on a risk-adjusted basis. Compounded until retirement, these losses can have a large impact on the welfare of retirees.

Our results in Table 11 reveal that plan participants do not benefit from the private information trustees may have about their own proprietary funds: trustee choices in 401(k) plans are not information driven. Instead, consistent with Carhart (1997), poor performance persists. Overall, plan participants invested into affiliated funds favored by trustees would have obtained a higher risk-adjusted performance had they switched their retirement savings from the underperforming trustee funds to other trustee funds.

7 Conclusion

Mutual fund families serving as trustees of 401(k) plans have a fiduciary duty to act in the interest of participants but they also have a competing incentive to attract and retain retirement contributions into their own proprietary funds. In this paper, we examine how trustee incentives influence the set of investment choices offered in the plan.

Despite the increasing importance of 401(k) plans as a retirement vehicle, little research has evaluated the consequence of offering the trusteeship of the plan to a mutual fund company. This is surprising, provided that small inefficiencies in the selection of investment options, especially early in the participant's career, can have a significant impact on retirement income.

Our paper takes a first step in this direction. We find that mutual fund trustees display favoritism toward their own funds. In particular, we show that trustee funds are less likely to be removed from the menu relative to non-trustee funds, independent of their performance record. Moreover, the difference in deletion propensities between trustee and non-trustee funds is largest among the worst performing funds. We find similar results for mutual fund additions.

Interestingly, mutual fund affiliation does not affect how participants allocate their contributions, suggesting that participants do not understand these biases. We also show that trustees' resistance to remove their own poorly performing funds generates a significant subsequent negative abnormal return of 2.9–3.6% per year for participants investing in those funds.

One question that we do not address in this paper is whether or not sponsoring companies should employ mutual fund families as plan trustees. Instead, we take a first step to uncover the various incentives that accompany the trustee relation and their effect on plan design.

Future research should explore and contrast additional costs and benefits of employing mutual fund and non-mutual fund trustees in the management of 401(k) plans.

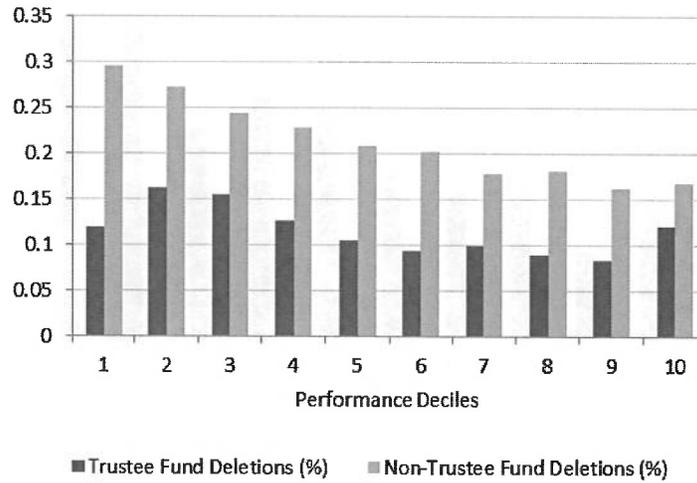
References

- Agnew, J., P. Balduzzi, and A. Sunden (2003). Portfolio choice and trading in a large 401(k) plan. *American Economic Review* 93, 193–215.
- Ai, C. and E. C. Norton (2003). Interaction terms in logit and probit models. *Economics Letters* 80, 123–129.
- Balduzzi, P. and J. Reuter (2012). Heterogeneity in target-date funds and the pension protection act of 2006. *Working paper*.
- Benartzi, S. and R. H. Thaler (2001). Naive diversification strategies in defined contribution saving plans. *American Economic Review* 91(1), 79–98.
- Bhattacharya, U., J. H. Lee, and V. K. Pool (2012). Conflicting family values in mutual fund families. *Forthcoming: Journal of Finance*.
- Brown, J. R., N. Liang, and S. Weisbenner (2007). Individual account investment options and portfolio choice: Behavioral lessons from 401(k) plans. *Journal of Public Economics* 91, 1992–2013.
- Brown, K. C. and W. V. Harlow (2012). How good are the investment options provided by defined contribution plan sponsors? *International Journal of Portfolio Analysis and Management* 1, 3–31.
- Carhart, M. M. (1997). On the persistence of mutual fund performance. *Journal of Finance* 52(1), 57–82.
- Carroll, G. D., J. J. Choi, D. Laibson, B. C. Madrian, and A. Metrick (2009). Optimal defaults and active decisions. *Quarterly Journal of Economics* 124, 1639–1674.
- Chaudhuri, R., Z. Ivkovic, and C. Trzcinka (2012). Strategic performance allocation in institutional asset management firms: Behold the power of stars and dominant clients. *Working paper*.
- Chevalier, J. A. and G. D. Ellison (1997). Risk taking in mutual funds as a response to incentives. *Journal of Political Economy* 105(6), 1167–1200.
- Choi, J. J., D. Laibson, B. C. Madrian, and A. Metrick (2002). Defined contribution pensions: Plan rules, participant decisions, and the path of least resistance. In J. M. Poterba (Ed.), *Tax Policy and the Economy*, pp. 67–113. Cambridge, MA: MIT Press.
- Choi, J. J., D. Laibson, B. C. Madrian, and A. Metrick (2004). For better or for worse. Default effects and 401(k) savings behavior. In D. A. Wise (Ed.), *Perspectives on the Economics of Aging*, pp. 81–121. Chicago, IL: University of Chicago Press.
- Christoffersen, S. K. and M. Simutin (2012). Risk-taking and retirement investing in mutual funds. *Working Paper*.
- Cohen, L. and B. Schmidt (2009). Attracting flows by attracting big clients. *The Journal of Finance* 64(5), 2125–2151.
- Davis, G. F. and E. H. Kim (2007). Business ties and proxy voting by mutual funds. *Journal of Financial Economics* 85(2), 552–570.

- Del Guercio, D. and P. Tkac (2002). The determinants of the flow of funds of managed portfolios: Mutual funds versus pension funds. *Journal of Financial and Quantitative Analysis* 37, 523–557.
- Duan, Y., E. S. Hotchkiss, and Y. Jiao (2012). Business ties and information advantage: Evidence from mutual fund trading. *Working Paper*.
- Duflo, E. and E. Saez (2002). Participation and investment decisions in a retirement plan: The influence of colleagues' choices. *Journal of Public Economics* 85(1), 121–148.
- Elton, E. J., M. J. Gruber, and C. R. Blake (2006). The adequacy of investment choices offered by 401(k) plans. *Journal of Public Economics* 90, 1299–1314.
- Elton, E. J., M. J. Gruber, and C. R. Blake (2007). Participant reaction and the performance of funds offered by 401(k) plans. *Journal of Financial Intermediation* 16, 249–271.
- Evans, R. B. (2010). Mutual fund incubation. *Journal of Finance* 65, 1581–1611.
- Fama, E. F. and K. R. French (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33(1), 3–56.
- Gaspar, J. M., M. Massa, and P. Matos (2006). Favoritism in mutual fund families? Evidence on strategic cross-fund subsidization. *Journal of Finance* 61, 73–104.
- Goyal, A. and S. Wahal (2008). The selection and termination of investment management firms by plan sponsors. *Journal of Finance* 63, 1805–1847.
- Huang, J., K. D. Wei, and H. Yan (2007). Participation costs and the sensitivity of fund flows to past performance. *Journal of Finance* 62, 1273–1311.
- Huberman, G. and W. Jiang (2006). Offering versus choice in 401(k) plans: Equity exposure and number of funds. *Journal of Finance* 61, 763–801.
- Kuhnen, C. M. (2009). Business networks, corporate governance, and contracting in the mutual fund industry. *Journal of Finance* 64, 2185–2220.
- Madrian, B. C. and D. F. Shea (2001). The power of suggestion: Inertia in 401(k) participation and savings behavior. *Quarterly Journal of Economics* 116, 1149–1187.
- Massa, M. and Z. Rehman (2008). Information flows within financial conglomerates: Evidence from the banks mutual funds relation. *Journal of Financial Economics* 89, 288–306.
- Mitchell, O. S. and S. Utkus (2012). Target-date funds in 401(k) retirement plans. *Working Paper*.
- Nanda, V., Z. J. Wang, and L. Zheng (2004). Family values and the star phenomenon: Strategies of mutual fund families. *Review of Financial Studies* 17(3), 667–698.
- Rauh, J. (2006). Own company stock in defined contribution pension plans: A takeover defense? *Journal of Financial Economics* 81(2), 379–410.
- Reuter, J. (2006). Are IPO allocations for sale? Evidence from mutual funds. *Journal of Finance* 61, 2289–2324.

- Ritter, J. R. and D. Zhang (2007). Affiliated mutual funds and the allocation of initial public offerings. *Journal of Financial Economics* 86, 337–368.
- Ruiz-Zaiko, L. and B. Williams (2007). Plan sponsors besieged by 401(k) fee lawsuits. *Pensions & Benefits Management Bridgebay Financial*.
- Sialm, C. and L. Starks (2012). Mutual fund tax clienteles. *Journal of Finance* 67, 1397–1422.
- Sialm, C., L. Starks, and H. Zhang (2012). Defined contribution pension plans: Sticky or discerning money? *Working Paper*.
- Sirri, E. R. and P. Tufano (1998). Costly search and mutual fund flows. *Journal of Finance* 53(5), 1598–1622.
- Tang, N., O. S. Mitchell, G. R. Mottola, and S. P. Utkus (2010). The efficiency of sponsor and participant portfolio choices in 401(k) plans. *Journal of Public Economics* 94, 1073–1085.

Panel A: Overall Sample



Panel B: Subsample of Funds on Both Affiliated and Unaffiliated Menus

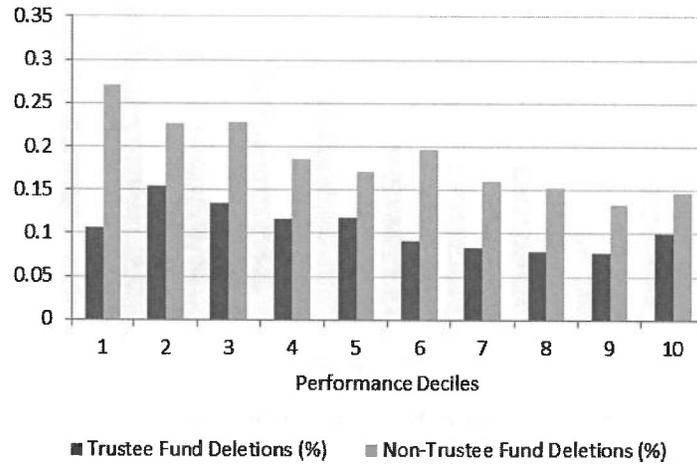
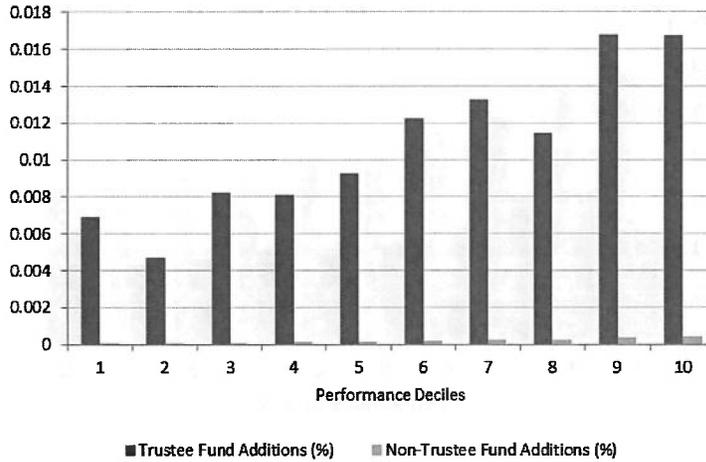


Figure 1: Fraction of Fund Deletions by Trustee and Non-Trustee Funds. The figure depicts the mean annual fund deletion frequencies by trustee affiliation and performance decile. Panel A includes the overall sample of mutual funds. Panel B includes the subsample of funds, where funds appear contemporaneously on multiple 401(k) menus, at least once as a trustee fund and at least once as a non-trustee fund. Every year, we calculate the ratio of the number of affiliated (unaffiliated) menus from which the fund is delisted during the year to the total number of affiliated (unaffiliated) menus associated with the fund. Performance is ranked using style-adjusted returns over the prior three years relative to the universe of mutual funds in CRSP. We then average across the funds' deletion frequencies by performance and affiliation.

Panel A: Overall Sample



Panel B: Subsample of Funds on Both Affiliated and Unaffiliated Menus

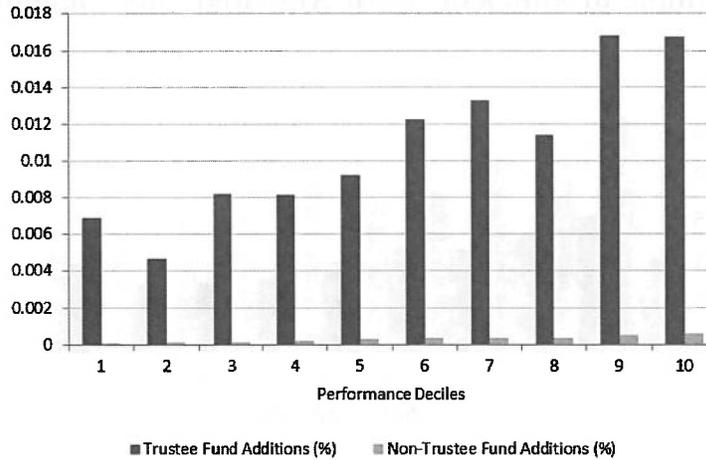


Figure 2: Fraction of Fund Additions by Mutual Fund and Non-Mutual Fund Trustees. The figure depicts the mean annual fund addition frequencies by trustee affiliation and performance decile. Panel A includes the overall sample of mutual funds. Panel B includes the subsample of funds, which are offered by fund families that serve as trustees for some firms in our sample. Every year, we calculate the ratio of the number of affiliated (unaffiliated) menus to which the fund is added during the year to the total number of affiliated (unaffiliated) menus that do not yet include the fund as an option. Performance is ranked using style-adjusted returns over the prior three years relative to the universe of mutual funds in CRSP. We then average across the funds' addition frequencies by performance and affiliation.

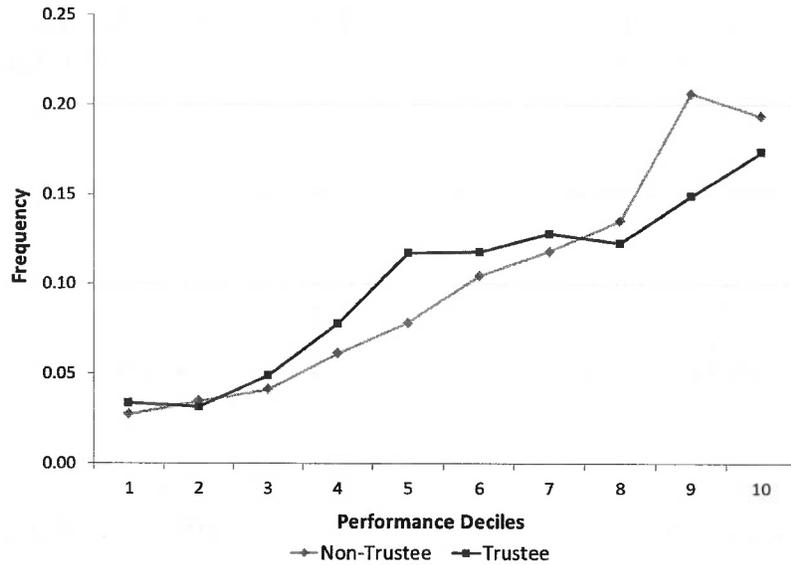


Figure 3: The Distribution of Mutual Funds Additions by Performance Decile and Fund Affiliation.

The figure shows the distribution of the funds that are added to a 401(k) menu at some point during our sample period by performance decile and affiliation. The dark line shows the fractions of trustee funds in the various performance deciles, while the grey line provides the corresponding values for non-trustee funds. Performance deciles are created from percentile performance ranks. These are calculated using overall overall rankings, in which fund performance is ranked relative to all other mutual funds with the same investment style in CRSP, based on returns in the prior 36 months.

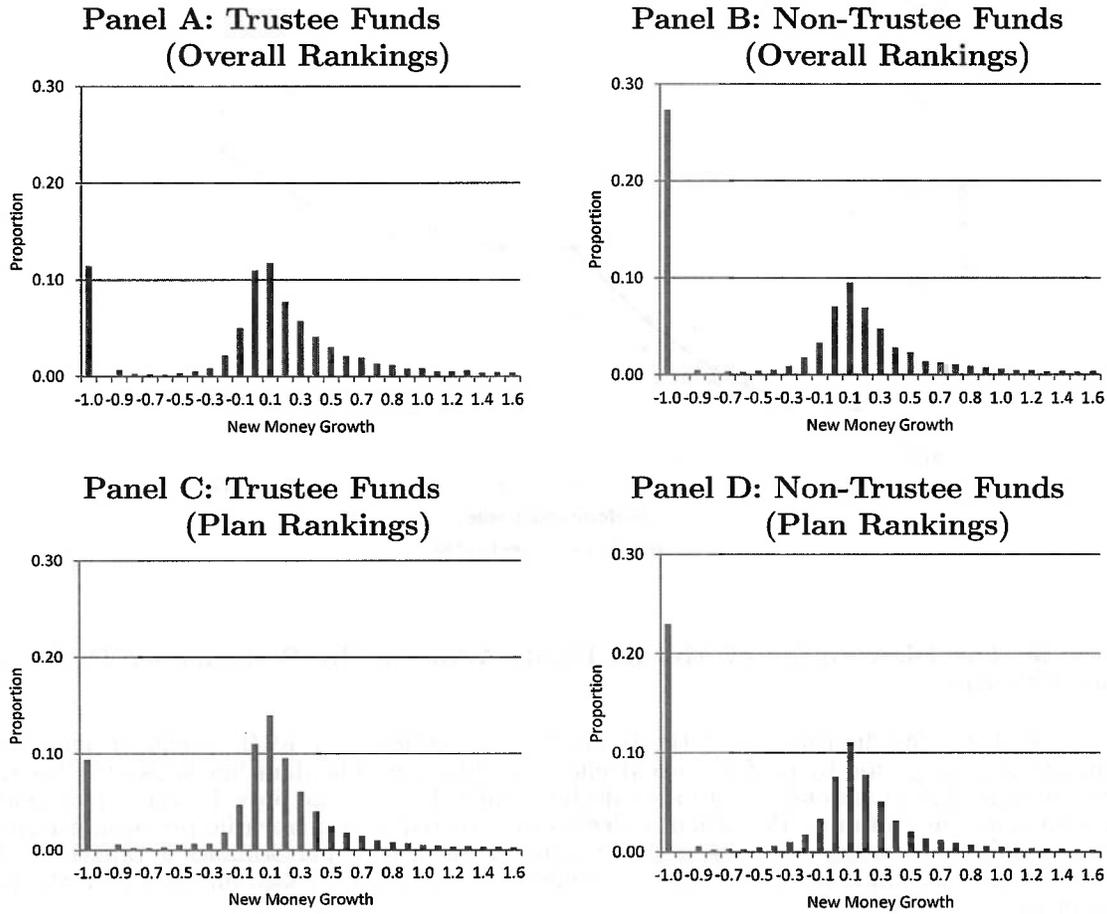


Figure 4: New Money Growth of Lower Performance Quintiles for Mutual Fund and Non-Mutual Fund Trustee Funds.

The figure displays the distribution of fund flows to poorly performing mutual funds on the menu by affiliation. Fund flows, or the growth rate of new money $NMG_{p,f,t}$ of fund f held in 401(k) plan p at time t is defined by $NMG_{p,f,t} = [V_{p,f,t} - V_{p,f,t-1}(1 + R_{f,t})] / [V_{p,f,t-1}(1 + R_{f,t})]$. The numerator captures the dollar change in the value of participants' investments ($V_{p,f,t}$) in fund f in plan p in year t after adjusting for the price appreciation $R_{f,t}$ (i.e., fund return) during the year. The denominator is defined as the projected value of the lagged plan position in the fund without any new flow of money. If an investment option is deleted from a plan menu, then NMG equals exactly -100%. In Panel A and B, the distributions describe fund flows to those trustee and non-trustee funds, respectively, that fall into the worst performance decile of the universe of mutual funds in their style category. Panels C and D depict the distributions of the corresponding flows using performance rankings based on only those mutual funds that are offered on the same 401(k) menu.

Table 1: Trustee Summary Statistics.

The table provides descriptive statistics by year. Columns 1 and 2 report the number of plans and plan sponsors captured in our sample, respectively. Columns 3 and 4 show the total number of mutual fund and non-mutual fund trustees, respectively, while columns 5 and 6 report the corresponding distinct number of trustees by type. In columns 7-11, we provide information about the architecture of the plan. These include the number of mutual fund options offered, the trustee share calculated as the overall proportion of retirement assets invested with affiliated funds, the average number of management companies that offer at least one investment option on the menu, and the Herfindahl index of the menu calculated based on the dollar share of each of these management companies.

Year	Total Number			Distinct Number			Plan Architecture				
	Sponsors	Plans	Mutual Fund Trustees	Other Trustees	Mutual Fund Trustees	Other Trustees	No. of Options	No. of Trustee Options	Trustee Share	No. of Mgmt. Comp.	Herfindahl Index
1998	621	735	439	296	73	57	6.98	2.35	0.34	2.96	0.67
1999	766	924	637	287	90	83	7.83	2.85	0.34	3.46	0.64
2000	835	1,035	758	277	89	95	9.25	3.54	0.36	3.98	0.59
2001	923	1,125	838	287	91	100	10.39	4.13	0.38	4.45	0.57
2002	1,013	1,250	955	295	97	107	11.45	4.63	0.38	4.97	0.55
2003	1,104	1,341	1,116	225	101	113	12.00	4.75	0.36	5.33	0.51
2004	1,105	1,333	1,112	221	100	112	13.16	5.23	0.34	5.76	0.48
2005	1,095	1,298	1,086	212	93	101	13.79	5.42	0.33	6.06	0.45
2006	1,034	1,238	968	270	93	93	14.58	5.83	0.32	6.25	0.44
2007	1,000	1,183	891	292	86	96	15.96	5.95	0.29	6.65	0.42
2008	970	1,135	859	276	82	82	17.23	6.53	0.29	7.06	0.42
2009	849	988	743	245	70	73	17.85	6.43	0.27	7.41	0.40
Total	11,315	13,585	10,402	3,183	1,065	1,112	13.11	5.07	0.33	5.62	0.494

Table 2: Mutual Funds Summary Statistics.

Panels A, B, and C of the table describe the funds that are deleted from, kept, or added to a 401(k) menu in our sample. We report the average fund age, fund size (in millions), the volatility of monthly fund returns, turnover, the raw and style-adjusted expense ratio, and the funds' mean percentile performance ranks. Performance ranks are calculated over the previous one, three, or five years based on overall rankings. The overall performance rank of each fund represents the performance of the fund relative to other funds in the same objective code. With the exception of fund age and fund size, all values are expressed as percentages. The averages are reported for trustee and non-trustee funds separately, along with the results of the paired *t*-test testing whether the difference is significantly different from zero. Standard errors for this test are two-way clustered at the plan and fund levels. Significance levels are denoted by *, **, ***, which correspond to 10%, 5%, and 1% levels, respectively.

Panel A: Mutual Funds Deletions

Trustee Fund	N	Fund Age	Fund Size	Return Std. Dev.	Turnover	Exp. Ratio	Exp. Ratio Style Adj.	Perf Rank 1 Yr.	Perf Rank 3 Yr.	Perf Rank 5 Yr.
0	15,380	17.98	7,810.90	4.08	92.49	1.06	0.17	48.00	50.67	55.13
1	3,623	17.55	7,249.91	3.43	78.86	0.79	-0.03	50.61	51.39	52.79
Diff		-0.42	-560.99	-0.65***	-13.63***	-0.27***	-0.20***	2.61**	0.72	-2.34**

Panel B: Mutual Funds Kept

Trustee Fund	N	Fund Age	Fund Size	Return Std. Dev.	Turnover	Exp. Ratio	Exp. Ratio Style Adj.	Perf Rank 1 Yr.	Perf Rank 3 Yr.	Perf Rank 5 Yr.
0	86,522	19.39	15,323.91	4.00	77.87	0.94	0.06	54.89	60.05	63.42
1	52,660	17.35	12,176.73	3.37	51.93	0.56	-0.22	55.44	58.07	59.28
Diff		-2.03	-3,147.18***	-0.62***	-25.94***	-0.38***	-0.29***	0.55	-1.98***	-4.15***

Panel C: Mutual Funds Additions

Trustee Fund	N	Fund Age	Fund Size	Return Std. Dev.	Turnover	Exp. Ratio	Exp. Ratio Style Adj.	Perf Rank 1 Yr.	Perf Rank 3 Yr.	Perf Rank 5 Yr.
0	21,196	15.10	9,962.67	3.99	83.80	0.96	0.07	62.31	67.39	68.61
1	6,997	10.32	5,443.62	3.23	55.78	0.59	-0.22	58.69	63.78	65.11
Diff		-4.78***	-4,519.06*	-0.77***	-28.02***	-0.36***	-0.28***	-3.62***	-3.62***	-3.50***

Table 3: Fund Deletion Proportions by Performance Deciles.

The table summarizes the mean annual fund deletion frequencies (as a %) by trustee affiliation and performance decile. Panel A includes the overall sample of mutual funds. Panel B includes the subsample of funds, where funds appear contemporaneously on multiple 401(k) menus, at least once as a trustee fund and at least once as a non-trustee fund. Every year, we calculate the ratio of the number of affiliated (unaffiliated) menus from which the fund is delisted during the year to the total number of affiliated (unaffiliated) menus associated with the fund. Performance is ranked using style-adjusted returns over the prior three years relative to the universe of mutual funds in CRSP. We then average across the funds' deletion frequencies by performance and affiliation. Significance levels are denoted by *, **, ***, which correspond to 10%, 5%, and 1% levels, respectively and are based on standard errors that are clustered at the fund level.

Panel A: All Funds

Performance Decile	1 Year			3 Years			5 Years		
	T	NT	T-NT	T	NT	T-NT	T	NT	T-NT
1	10.8	25.5	-14.7***	11.9	29.6	-17.7***	13.1	27.2	-14.2***
2	15.2	23.7	-8.6***	16.2	27.2	-11.0***	15.3	27.9	-12.6***
3	13.1	23.1	-10.0***	15.5	24.5	-9.0***	13.8	24.8	-10.9***
4	12.9	23.4	-10.5***	12.6	22.9	-10.3***	11.3	24.3	-13.0***
5	11.8	19.6	-7.7***	10.6	20.8	-10.2***	11.7	21.2	-9.5***
6	11.1	19.1	-7.9***	9.3	20.3	-10.9***	10.7	20.9	-10.2***
7	8.5	19.4	-10.9***	10.0	17.8	-7.9***	9.8	19.8	-10.1***
8	10.5	18.7	-8.2***	9.0	18.1	-9.1***	9.3	17.0	-7.7***
9	8.5	18.0	-9.5***	8.3	16.2	-7.9***	9.3	15.9	-6.7***
10	10.9	17.3	-6.5***	12.2	16.9	-4.7***	11.3	17.0	-5.8***

Panel B: Subsample of Funds on Both Affiliated and Unaffiliated Menus

Performance Decile	1 Year			3 Years			5 Years		
	T	NT	T-NT	T	NT	T-NT	T	NT	T-NT
1	9.1	21.3	-12.2***	10.5	27.0	-16.5***	13.5	25.9	-12.4***
2	13.5	21.7	-8.2***	15.4	22.6	-7.2***	13.4	25.6	-12.3***
3	13.4	20.4	-7.0***	13.5	22.7	-9.2***	11.8	21.5	-9.7***
4	11.5	20.9	-9.4***	11.6	18.6	-7.1***	11.6	19.3	-7.7***
5	12.4	19.0	-6.6***	11.8	17.0	-5.2***	11.0	18.9	-7.9***
6	9.1	15.8	-6.7***	9.0	19.6	-10.6***	11.0	18.6	-7.6***
7	6.9	14.7	-7.8***	8.3	16.1	-7.8***	9.0	18.3	-9.3***
8	9.5	15.5	-6.0***	7.9	15.3	-7.4***	7.9	15.5	-7.5***
9	8.3	16.7	-8.4***	7.8	13.3	-5.5***	7.9	11.9	-4.0***
10	9.3	14.5	-5.3***	10.0	14.7	-4.6**	9.7	13.7	-4.0**

Table 4: Linear Probability Model for Fund Deletions.

The table reports the OLS coefficient estimates of the following model for fund deletions: $DEL_{p,f,t} = \beta_0 + \beta_1 \times TF_{p,f,t} + \beta_2 \times LowRank_{p,f,t} + \beta_3 \times HighRank_{p,f,t} + \beta_4 \times TF_{p,f,t} \times LowRank_{p,f,t} + \beta_5 \times TF_{p,f,t} \times HighRank_{p,f,t} + Z'_{p,f,t} \gamma + \epsilon_{p,f,t}$, where $DEL_{p,f,t}$ is an indicator variable that takes the value of one if mutual fund f has been deleted from plan p at time t and zero otherwise, and $TF_{p,f,t}$ is an indicator variable for whether the trustee of pension plan p is affiliated with the management company of mutual fund f . $LowRank$ and $HighRank$ are defined as $LowRank_{p,f,t} = \min(Rank_{p,f,t}, 0.5)$ and $HighRank_{p,f,t} = \min(Rank_{p,f,t} - LowRank_{p,f,t}, 0.5)$, where $Rank_{p,f,t}$ is the percentile performance rank of mutual fund f over the previous one, three, or five years based on the performance of the fund relative to other CRSP funds in the same objective code. The other control variables Z include the natural logarithm of the option size, the number of options, the expense ratio of the fund, the turnover of the fund's holdings, the natural logarithm of the fund's size, fund age, the standard deviation of the fund's return, and unreported indicator variables for specific fund types and year fixed effects. Standard errors are two-way clustered at the plan and fund levels and are reported in parentheses. Significance levels are denoted by *, **, ***, which correspond to 10%, 5%, and 1% levels, respectively.

	1 Year	3 Years	5 Years
Trustee Fund	-0.099*** (0.015)	-0.140*** (0.018)	-0.119*** (0.022)
LowRank	-0.181*** (0.029)	-0.324*** (0.034)	-0.230*** (0.037)
HighRank	-0.054** (0.024)	-0.072*** (0.023)	-0.164*** (0.024)
LowRank*Trustee Fund	0.171*** (0.035)	0.247*** (0.042)	0.152*** (0.052)
HighRank*Trustee Fund	-0.020 (0.030)	-0.003 (0.027)	0.085*** (0.030)
Log(Option Size)	-0.007*** (0.002)	-0.007*** (0.002)	-0.008*** (0.002)
No. of Options	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)
Exp. Ratio	5.070*** (0.948)	4.611*** (0.931)	5.169*** (0.976)
Turnover	0.013*** (0.004)	0.013*** (0.004)	0.014*** (0.004)
Log(Fund Size)	-0.023*** (0.002)	-0.021*** (0.002)	-0.019*** (0.002)
Fund Age	0.000** (0.000)	0.000 (0.000)	0.000 (0.000)
Std. Dev.	-0.046 (0.207)	0.262 (0.207)	0.370* (0.206)
Observations	99,967	99,967	99,967
R-Squared	0.061	0.069	0.066

Table 5: Linear Probability Model for Fund Deletions: Alternative Rankings.

The table reports the OLS coefficient estimates of the following model for fund deletions: $DEL_{p,f,t} = \beta_0 + \beta_1 \times TF_{p,f,t} + \beta_2 \times LowRank_{p,f,t} + \beta_3 \times HighRank_{p,f,t} + \beta_4 \times TF_{p,f,t} \times LowRank_{p,f,t} + \beta_5 \times TF_{p,f,t} \times HighRank_{p,f,t} + Z'_{p,f,t} \gamma + \epsilon_{p,f,t}$, where $DEL_{p,f,t}$ is an indicator variable that takes the value of one if mutual fund f has been deleted from plan p at time t and zero otherwise, and $TF_{p,f,t}$ is an indicator variable for whether the trustee of pension plan p is affiliated with the management company of mutual fund f . $LowRank$ and $HighRank$ are defined as $LowRank_{p,f,t} = \min(Rank_{p,f,t}, 0.5)$ and $HighRank_{p,f,t} = \min(Rank_{p,f,t} - LowRank_{p,f,t}, 0.5)$, where $Rank_{p,f,t}$ is the percentile performance rank of mutual fund f over the previous one, three, or five years based either on the fund's percentile rankings within a specific 401(k) plan or on the fund's percentile rankings within the fund's family. The other control variables Z include the natural logarithm of the option size, the number of options, the expense ratio of the fund, the turnover of the fund's holdings, the natural logarithm of the fund's size, fund age, the standard deviation of the fund's return, and unreported indicator variables for specific fund types and year fixed effects. Standard errors are two-way clustered at the plan and fund levels and are reported in parentheses. Significance levels are denoted by *, **, ***, which correspond to 10%, 5%, and 1% levels, respectively.

	Plan Ranking			Family Ranking		
	1 Year	3 Years	5 Years	1 Year	3 Years	5 Years
Trustee Fund	-0.081*** (0.012)	-0.088*** (0.014)	-0.103*** (0.015)	-0.070*** (0.014)	-0.075*** (0.016)	-0.076*** (0.017)
LowRank	-0.135*** (0.020)	-0.231*** (0.024)	-0.185*** (0.020)	-0.073*** (0.024)	-0.164*** (0.029)	-0.159*** (0.033)
HighRank	-0.042* (0.023)	-0.055** (0.022)	-0.186*** (0.019)	-0.052** (0.025)	-0.052** (0.025)	-0.078*** (0.027)
LowRank*Trustee Fund	0.134*** (0.030)	0.139*** (0.031)	0.110*** (0.024)	0.099*** (0.033)	0.116*** (0.041)	0.106** (0.044)
HighRank*Trustee Fund	-0.015 (0.030)	0.019 (0.030)	0.105*** (0.024)	-0.020 (0.033)	-0.038 (0.036)	-0.013 (0.036)
Log(Option Size)	-0.007*** (0.002)	-0.007*** (0.002)	-0.008*** (0.002)	-0.007*** (0.002)	-0.007*** (0.002)	-0.007*** (0.002)
No. of Options	-0.002*** (0.000)	-0.002*** (0.000)	-0.001*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)
Exp. Ratio	5.553*** (0.953)	5.465*** (0.938)	5.246*** (0.971)	5.439*** (0.960)	5.424*** (0.948)	5.619*** (0.943)
Turnover	0.012*** (0.003)	0.011*** (0.003)	0.014*** (0.004)	0.013*** (0.004)	0.013*** (0.004)	0.013*** (0.003)
Log(Fund Size)	-0.024*** (0.002)	-0.022*** (0.002)	-0.020*** (0.002)	-0.024*** (0.002)	-0.024*** (0.002)	-0.023*** (0.002)
Fund Age	0.000** (0.000)	0.000 (0.000)	0.000 (0.000)	0.000** (0.000)	0.000** (0.000)	0.000* (0.000)
Std. Dev.	-0.116 (0.213)	0.131 (0.215)	0.383* (0.211)	-0.003 (0.217)	0.106 (0.208)	0.264 (0.205)
Observations	100,299	100,299	100,299	100,269	100,269	100,269
R-squared	0.059	0.065	0.066	0.057	0.059	0.060

Table 6: Linear Probability Model for Fund Deletions: Alternative Functional Forms. The table reports the OLS coefficient estimates of the model for fund deletions described in Table 4 using a linear performance specification (one-segment model) in columns 1-3, and a three-segment piecewise linear specification in columns 4-6. For the three-segment specification, the performance segments are 1) the lowest performance quintile, 2) the highest performance quintile, and 3) the three middle performance quintiles, which are pulled together to represent a single performance segment. Following Sirri and Tufano (1998), the performance in the lowest quintile is given by $LowQRank_{p,f,t} = \min(Rank_{p,f,t}, 0.2)$, the performance in the three middle quintiles is given by $MidQRank_{p,f,t} = \min(Rank_{p,f,t} - LowQRank_{p,f,t}, 0.6)$, and the performance in the highest quintile is given by $HighQRank_{p,f,t} = (Rank_{p,f,t} - LowQRank_{p,f,t} - MidQRank_{p,f,t})$, where $Rank_{p,f,t}$ is the percentage performance rank of mutual fund f over the previous three years based on either overall rankings, 401(k) plan rankings, and fund family rankings. Standard errors in this table are two-way clustered at the plan and fund levels and are reported in parentheses. Significance levels are denoted by *, **, ***, which correspond to 10%, 5%, and 1% levels, respectively.

Ranking	Linear Performance			Three Segments		
	Overall	Plan	Family	Overall	Plan	Family
Trustee Fund	-0.095*** (0.012)	-0.071*** (0.012)	-0.051*** (0.012)	-0.164*** (0.031)	-0.106*** (0.025)	-0.105*** (0.027)
Rank	-0.181*** (0.015)	-0.145*** (0.017)	-0.103*** (0.015)			
LowQRank				-0.467*** (0.120)	-0.576*** (0.094)	-0.407*** (0.112)
MidQRank				-0.210*** (0.019)	-0.111*** (0.015)	-0.082*** (0.020)
HighQRank				0.215*** (0.074)	-0.026 (0.095)	-0.068 (0.077)
Rank*Trustee Fund	0.103*** (0.017)	0.079*** (0.019)	0.034* (0.018)			
LowQRank*Trustee Fund				0.479*** (0.164)	0.293** (0.126)	0.359** (0.145)
MidQRank*Trustee Fund				0.112*** (0.023)	0.064*** (0.020)	0.012 (0.025)
HighQRank*Trustee Fund				-0.176* (0.095)	0.001 (0.123)	-0.009 (0.122)
Log(Option Size)	-0.007*** (0.002)	-0.007*** (0.002)	-0.007*** (0.002)	-0.007*** (0.002)	-0.007*** (0.002)	-0.007*** (0.002)
No. of Options	-0.001*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)	-0.001*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)
Exp. Ratio	4.884*** (0.941)	5.524*** (0.940)	5.485*** (0.949)	4.539*** (0.934)	5.470*** (0.937)	5.376*** (0.947)
Turnover	0.013*** (0.004)	0.012*** (0.003)	0.013*** (0.004)	0.014*** (0.004)	0.011*** (0.003)	0.013*** (0.004)
Log(Fund Size)	-0.021*** (0.002)	-0.022*** (0.002)	-0.024*** (0.002)	-0.021*** (0.002)	-0.022*** (0.002)	-0.024*** (0.002)
Fund Age	0.000 (0.000)	0.000 (0.000)	0.000* (0.000)	0.000 (0.000)	0.000 (0.000)	0.000* (0.000)
Std. Dev.	0.326 (0.213)	0.180 (0.216)	0.126 (0.206)	0.247 (0.204)	0.111 (0.216)	0.107 (0.209)
Observations	99,967	100,299	100,269	99,967	100,299	100,269
R-squared	0.067	0.064	0.059	0.069	0.065	0.060

Table 7: Linear Probability Model for Fund Deletions: Robustness Tests.

The table reports the OLS coefficient estimates of the model for fund deletions described in Table 4 for various subsamples of our data. We exclude the three largest trustees each year in the first column, estimate the model for the three largest trustees in the second, and estimate our model with trustee fixed effects in column 3. Column 4 reestimates our results using information only on those plans that are trustee by a mutual fund family. Column 5 excludes target date funds and column 6 excludes all non-equity funds. Finally, in columns 7 and 8, we divide our sample into the subperiods 1998-2006 and 2007-2009, respectively. Standard errors in this table are two-way clustered at the plan and fund levels and are reported in parentheses. Significance levels are denoted by *, **, ***, which correspond to 10%, 5%, and 1% levels, respectively.

	Exclude Top 3 MF Trustees	Exclude Top 3 MF Trustees	Include Trustee FE	Only MF Trustees	Exclude Target Date Funds	Only Equity Funds	Prior to 2007	After 2006
Trustee Fund	-0.139*** (0.022)	-0.127*** (0.037)	-0.157*** (0.017)	-0.149*** (0.020)	-0.131*** (0.020)	-0.117*** (0.023)	-0.117*** (0.029)	-0.147*** (0.022)
LowRank	-0.316*** (0.034)	-0.312*** (0.074)	-0.320*** (0.034)	-0.329*** (0.039)	-0.366*** (0.035)	-0.385*** (0.038)	-0.305*** (0.051)	-0.334*** (0.041)
HighRank	-0.068*** (0.023)	-0.128*** (0.045)	-0.076*** (0.021)	-0.065** (0.027)	-0.083*** (0.024)	-0.098*** (0.024)	-0.125*** (0.030)	-0.020 (0.030)
LowRank*Trustee Fund	0.178*** (0.055)	0.247*** (0.082)	0.242*** (0.040)	0.254*** (0.047)	0.208*** (0.045)	0.181*** (0.052)	0.236*** (0.065)	0.228*** (0.051)
HighRank*Trustee Fund	0.009 (0.037)	0.036 (0.048)	0.008 (0.026)	-0.012 (0.032)	0.030 (0.031)	0.023 (0.033)	0.036 (0.038)	-0.037 (0.034)
Log(Option Size)	-0.008*** (0.002)	-0.006* (0.003)	-0.008*** (0.001)	-0.008*** (0.002)	-0.009*** (0.002)	-0.009*** (0.002)	-0.007*** (0.002)	-0.008*** (0.002)
No. of Options	-0.001*** (0.000)	-0.003*** (0.001)	-0.002*** (0.000)	-0.002*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.001)	-0.001*** (0.000)
Exp. Ratio	3.920*** (1.025)	9.453*** (1.963)	4.534*** (0.890)	4.870*** (1.057)	4.413*** (1.024)	3.343*** (1.114)	5.876*** (1.385)	4.303*** (1.088)
Turnover	0.011*** (0.003)	0.020*** (0.007)	0.011*** (0.003)	0.012*** (0.004)	0.011*** (0.003)	0.026*** (0.007)	0.015*** (0.006)	0.009** (0.004)
Log(Fund Size)	-0.023*** (0.002)	-0.023*** (0.004)	-0.024*** (0.002)	-0.020*** (0.002)	-0.021*** (0.002)	-0.020*** (0.002)	-0.025*** (0.003)	-0.019*** (0.002)
Fund Age	0.000** (0.000)	0.000 (0.000)	0.000* (0.000)	0.000 (0.000)	0.000* (0.000)	0.000 (0.000)	0.000 (0.000)	0.000 (0.000)
Std. Dev.	0.496** (0.228)	-0.603** (0.239)	0.188 (0.133)	0.263 (0.227)	0.474** (0.207)	0.600** (0.239)	-1.009*** (0.364)	0.579*** (0.225)
Observations	63,996	36,303	94,153	82,295	85,899	66,341	49,412	50,887
R-squared	0.059	0.091	0.093	0.073	0.065	0.076	0.075	0.068

Table 8: Fund Addition Proportions by Performance Deciles.

The table summarizes the mean annual fund addition frequencies (as a %) by trustee affiliation and performance decile. Panel A includes the overall sample of mutual funds. Panel B includes the subsample of funds, which are offered by fund families that serve as trustees for some firms in our sample. Every year, we calculate the ratio of the number of affiliated (unaffiliated) menus to which the fund is added during the year to the total number of affiliated (unaffiliated) menus that do not yet include the fund as an option at the beginning of the year. Performance is ranked using style-adjusted returns over the prior three years relative to the universe of mutual funds in CRSP. We then average across the funds' addition frequencies by performance and affiliation. Significance levels are based on standard errors that are clustered at the fund level and are denoted by *, **, ***, which correspond to 10%, 5%, and 1% levels, respectively.

Panel A: All Funds

Performance Decile	1 Year			3 Years			5 Years		
	T	NT	T-NT	T	NT	T-NT	T	NT	T-NT
1	1.038	0.008	1.030***	0.690	0.005	0.685***	0.573	0.005	0.569***
2	0.613	0.009	0.604***	0.470	0.007	0.463***	0.497	0.005	0.492***
3	0.930	0.012	0.919***	0.820	0.008	0.812***	0.789	0.009	0.780***
4	0.896	0.014	0.881***	0.813	0.011	0.801***	0.658	0.011	0.647***
5	0.949	0.017	0.932***	0.924	0.014	0.909***	0.863	0.014	0.849***
6	1.045	0.020	1.025***	1.223	0.020	1.203***	0.916	0.016	0.900***
7	1.042	0.020	1.022***	1.327	0.021	1.305***	1.424	0.020	1.404***
8	1.328	0.024	1.304***	1.142	0.024	1.118***	1.375	0.027	1.348***
9	1.530	0.030	1.500***	1.681	0.037	1.644***	1.861	0.040	1.820***
10	1.434	0.034	1.400***	1.676	0.039	1.637***	1.856	0.039	1.817***

Panel B: Subsample of Funds on Both Affiliated and Unaffiliated Menus

Performance Decile	1 Year			3 Years			5 Years		
	T	NT	T-NT	T	NT	T-NT	T	NT	T-NT
1	1.038	0.017	1.021***	0.690	0.009	0.681***	0.573	0.008	0.565***
2	0.613	0.011	0.601***	0.470	0.010	0.459***	0.497	0.006	0.490***
3	0.930	0.019	0.912***	0.820	0.013	0.807***	0.789	0.015	0.775***
4	0.896	0.023	0.873***	0.813	0.019	0.794***	0.658	0.021	0.638***
5	0.949	0.027	0.922***	0.924	0.029	0.895***	0.863	0.028	0.835***
6	1.045	0.034	1.011***	1.223	0.034	1.189***	0.916	0.030	0.887***
7	1.042	0.032	1.009***	1.327	0.037	1.290***	1.424	0.032	1.392***
8	1.328	0.039	1.289***	1.142	0.036	1.106***	1.375	0.036	1.339***
9	1.530	0.047	1.483***	1.681	0.051	1.630***	1.861	0.061	1.800***
10	1.434	0.044	1.390***	1.676	0.056	1.619***	1.856	0.060	1.796***

Table 9: Linear Probability Model for Trustee Additions.

The table reports the OLS coefficient estimates of the following model for trustee fund additions: $TF^{ADD}_{p,f,t} = \beta_0 + \beta_1 \times Rank_{p,f,t} + Z'_{p,f,t}\gamma + \epsilon_{p,f,t}$, where $TF^{ADD}_{p,f,t}$ is an indicator variable equal to one if mutual fund f added to the plan p at time t is affiliated with the management company acting as the plan's trustee and zero otherwise. $Rank_{p,f,t}$ is the performance rank of mutual fund f over the previous one, three, or five years based on overall rankings, and is included as a percentage. The overall performance rank of each fund depends on the performance of the fund relative to other funds in the same objective code. The other control variables Z include the number of options, the expense ratio of the fund, the turnover of the fund, the natural logarithm of the fund's size, fund age, the standard deviation of the fund's return (all measured during the previous year), and unreported indicator variables for specific fund styles, and year and trustee fixed effects. Standard errors are two-way clustered at plan and fund levels and are reported in parentheses. Significance levels are denoted by *, **, ***, which correspond to 10%, 5%, and 1% levels, respectively.

	1 Year	3 Years	5 Years
Rank (1 Yr.)	-0.136*** (0.028)		
Rank (3 Yrs.)		-0.191*** (0.037)	
Rank (5 Yrs.)			-0.226*** (0.043)
No. of Options	-0.001** (0.000)	-0.001** (0.000)	-0.001** (0.000)
Log(Plan Assets)	-0.016*** (0.003)	-0.016*** (0.003)	-0.016*** (0.003)
Exp. Ratio	-0.138*** (0.024)	-0.135*** (0.024)	-0.131*** (0.024)
Turnover	0.000 (0.010)	-0.001 (0.010)	-0.002 (0.010)
Log(Fund Size)	-0.006 (0.007)	-0.004 (0.007)	-0.000 (0.007)
Fund Age	-0.002** (0.001)	-0.002*** (0.001)	-0.002*** (0.001)
Std. Dev.	0.026 (0.033)	0.050 (0.033)	0.054* (0.033)
Observations	16,511	16,511	16,511
R-squared	0.741	0.742	0.743

Table 10: Fund Flow Regressions.

The table reports the coefficient estimates of the following OLS regression: $NMG_{p,f,t} = \beta_0 + \beta_1 \times TF_{p,f,t} + \beta_2 \times LowRank_{p,f,t} + \beta_3 \times HighRank_{p,f,t} + \beta_4 \times TF_{p,f,t} \times LowRank_{p,f,t} + \beta_5 \times TF_{p,f,t} \times HighRank_{p,f,t} + Z'_{p,f,t}\gamma + \epsilon_{p,f,t}$, where the explanatory variables of the regression are analogous to those in Table 4 with the exception of *Plan Growth*, which is a new variable added in this table. Our first dependent variable (with corresponding results reported in columns 1 and 4 for all flows and participants, respectively) is new money growth defined as $NMG1_{p,f,t} = \frac{V_{p,f,t} - V_{p,f,t-1}(1+R_{f,t})}{V_{p,f,t-1}(1+R_{f,t})}$, where $V_{p,f,t}$ is the value of participants' investments in fund f in plan p in year t and $R_{f,t}$ is the fund's return during the year. We use two additional definitions for new money growth. $NMG2$ is new money growth defined as $NMG2_{p,f,t} = \frac{V_{p,f,t} - V_{p,f,t-1}(1+R_{f,t})}{V_{p,f,t} + V_{p,f,t-1}(1+R_{f,t})}$, with corresponding results reported in columns 2 and 5. Finally, $NMG3$ shares the numerator with the previous two definitions but we replace the denominator by lagged plan size. Regression results using $NMG3$ as the dependent variable are reported in columns 3 and 6. The performance rank of mutual fund f is calculated over the previous three years based on the overall ranking. Standard errors in this table are two-way clustered at the plan and fund levels and are reported in parentheses. Significance levels are denoted by *, **, ***, which correspond to 10%, 5%, and 1% levels, respectively.

	All Fund Flows			Participant Flows Only		
	NMG1	NMG2	NMG3	NMG1	NMG2	NMG3
Trustee Fund	0.249*** (0.041)	0.178*** (0.024)	1.009*** (0.246)	0.064* (0.036)	0.022* (0.013)	0.004 (0.132)
LowRank	0.006*** (0.001)	0.004*** (0.000)	0.037*** (0.005)	0.002*** (0.001)	0.001*** (0.000)	0.012*** (0.003)
HighRank	0.004*** (0.001)	0.002*** (0.000)	0.018*** (0.003)	0.003*** (0.000)	0.001*** (0.000)	0.009*** (0.003)
LowRank*Trustee Fund	-0.005*** (0.001)	-0.003*** (0.001)	-0.012** (0.006)	-0.001 (0.001)	-0.000 (0.000)	-0.005 (0.003)
HighRank*Trustee Fund	0.001 (0.001)	0.000 (0.000)	-0.013*** (0.004)	0.000 (0.001)	-0.000 (0.000)	-0.002 (0.003)
Plan Growth	0.871*** (0.073)	0.394*** (0.041)	5.912*** (0.432)	0.886*** (0.069)	0.374*** (0.026)	7.669*** (0.397)
Log(Option Size)	-0.086*** (0.005)	-0.024*** (0.003)	-0.241*** (0.017)	-0.111*** (0.004)	-0.036*** (0.002)	-0.075*** (0.011)
No. Options	-0.001 (0.001)	0.000 (0.001)	0.009** (0.004)	-0.003** (0.001)	-0.001*** (0.000)	-0.020*** (0.003)
Exp. Ratio	-12.121*** (2.153)	-6.913*** (1.239)	-26.163** (11.041)	-6.923*** (1.790)	-2.228*** (0.617)	1.897 (6.032)
Turnover	-0.018** (0.007)	-0.015*** (0.005)	-0.077*** (0.025)	0.001 (0.006)	0.000 (0.002)	-0.002 (0.018)
Log(Fund Size)	0.027*** (0.005)	0.025*** (0.003)	0.232*** (0.029)	-0.001 (0.005)	0.001 (0.002)	0.083*** (0.017)
Fund Age	-0.001*** (0.000)	-0.001** (0.000)	-0.003 (0.003)	-0.001*** (0.000)	-0.000*** (0.000)	-0.001 (0.002)
Std. Dev.	-0.530 (0.486)	-0.371 (0.291)	5.085*** (1.756)	0.215 (0.348)	0.046 (0.123)	2.517** (1.218)
Observations	89,276	89,276	89,276	77,911	77,911	77,911
R-squared	0.168	0.105	0.056	0.251	0.222	0.108

Table 11: Abnormal Returns of Mutual Fund and Non-Mutual Fund Trustees.

Panels A, B, and C of the table report the abnormal return $\alpha_{f,t}$ of fund portfolio f at time t using the Fama-French-Carhart four-factor model (FFM), the Fama and French (1993) model, and the CAPM model, respectively, over our complete sample period using monthly fund return data. At the end of each calendar year, we form equal-weighted portfolios of trustee and non-trustee domestic equity funds separately based on whether the funds were kept, deleted, or added to the 401(k) menu (“No Changes,” “Deletions,” and “Additions”) during the calendar year. This creates six portfolios. We then further subdivide these six groups based on past performance. In particular, “All Funds,” refers to the original six portfolios and “Lowest Quintile,” (“Lowest Decile”) refers to a sub-portfolio in each group that contains only those funds that also rank in the lowest performance quintile (decile) based on past performance. The panels report results using the overall performance rankings during the prior three years. The performance measures are reported in % per month. Robust standard errors are reported in parentheses. Significance levels are denoted by *, **, ***, which correspond to 10%, 5%, and 1% levels, respectively.

Panel A: Carhart Alphas

	No Changes		Deletions		Additions	
	Trustee Funds	Non-Trustee Funds	Trustee Funds	Non-Trustee Funds	Trustee Funds	Non-Trustee Funds
All Funds	-0.04 (0.04)	-0.06 (0.04)	-0.08 (0.06)	-0.06 (0.04)	-0.04 (0.05)	-0.06 (0.04)
Lowest Quintile	-0.24** (0.10)	-0.09 (0.09)	-0.18 (0.13)	-0.12 (0.10)	-0.15 (0.15)	-0.01 (0.11)
Lowest Decile	-0.30** (0.13)	-0.07 (0.14)	-0.35** (0.17)	-0.16 (0.16)	0.03 (0.28)	0.08 (0.18)

Panel B: Fama-French Alphas

	No Changes		Deletions		Additions	
	Trustee Funds	Non-Trustee Funds	Trustee Funds	Non-Trustee Funds	Trustee Funds	Non-Trustee Funds
All Funds	-0.04 (0.04)	-0.06 (0.04)	-0.08 (0.06)	-0.06 (0.04)	-0.04 (0.05)	-0.06 (0.04)
Lowest Quintile	-0.24** (0.10)	-0.09 (0.10)	-0.18 (0.13)	-0.12 (0.11)	-0.15 (0.15)	-0.00 (0.11)
Lowest Decile	-0.30** (0.13)	-0.07 (0.16)	-0.35** (0.17)	-0.16 (0.17)	0.02 (0.28)	0.08 (0.19)

Panel C: CAPM Alphas

	No Changes		Deletions		Additions	
	Trustee Funds	Non-Trustee Funds	Trustee Funds	Non-Trustee Funds	Trustee Funds	Non-Trustee Funds
All Funds	0.02 (0.05)	0.02 (0.04)	-0.04 (0.06)	-0.00 (0.05)	-0.00 (0.05)	0.04 (0.04)
Lowest Quintile	0.00 (0.14)	0.17 (0.15)	-0.07 (0.14)	0.10 (0.16)	0.02 (0.16)	0.24 (0.15)
Lowest Decile	-0.07 (0.16)	0.24 (0.26)	-0.22 (0.21)	0.09 (0.29)	0.08 (0.31)	0.34 (0.24)

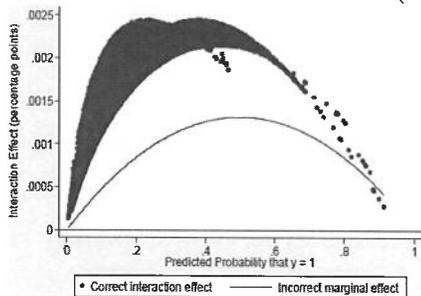
Appendix

Table A1: Probit Model for Fund Deletions: Two Piecewise Linear Segments.

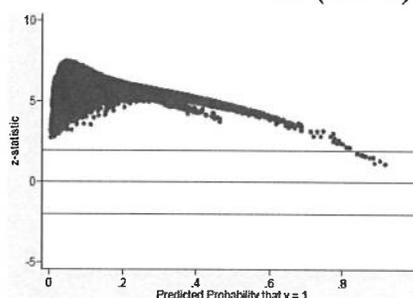
The table reports the estimated marginal effects of the following probit model for fund deletions: $DEL_{p,f,t} = \beta_0 + \beta_1 \times TF_{p,f,t} + \beta_2 \times LowRank_{p,f,t} + \beta_3 \times HighRank_{p,f,t} + \beta_4 \times TF_{p,f,t} \times HighRank_{p,f,t} + \beta_5 \times TF_{p,f,t} \times LowRank_{p,f,t} + Z'_{p,f,t} \gamma + \epsilon_{p,f,t}$, where $DEL_{p,f,t}$ is an indicator variable that takes the value of one if mutual fund f has been deleted from plan p at time t and zero otherwise, and $TF_{p,f,t}$ is an indicator variable for whether the trustee of pension plan p is affiliated with the management company of mutual fund f . $LowRank$ and $HighRank$ are defined as $LowRank_{p,f,t} = \min(Rank_{p,f,t}, 0.5)$ and $HighRank_{p,f,t} = \min(Rank_{p,f,t} - LowRank_{p,f,t}, 0.5)$, where $Rank_{p,f,t}$ is the performance rank of mutual fund f over the previous one, three, or five years based on overall rankings, and is included as a percentage. The overall performance rank of each fund depends on the performance of the fund relative to other funds in the same objective code, whereas the inside performance rank only depends on the fund's ranking within the 401(k) plan. The other control variables Z include the natural logarithm of the option size, the number of options, the expense ratio of the fund, the turnover of the fund's holdings, the natural logarithm of the fund's size, fund age, the standard deviation of the fund's return, and unreported indicator variables for specific fund types and year fixed effects. The marginal effects for the interaction terms are computed using the INTEFF command based on Standard errors are clustered at the plan level and are reported in parentheses. Significance levels are denoted by *, **, ***, which correspond to 10%, 5%, and 1% levels, respectively.

	Overall Ranking			Plan Ranking		
	1 Yr.	3 Yrs.	5 Yrs.	1 Yr.	3 Yrs.	5 Yrs.
Trustee Fund	-0.065*** (0.008)	-0.074*** (0.008)	-0.059*** (0.010)	-0.059*** (0.007)	-0.051*** (0.007)	-0.052*** (0.007)
LowRank	-0.129*** (0.011)	-0.208*** (0.013)	-0.130*** (0.014)	-0.093*** (0.010)	-0.155*** (0.010)	-0.134*** (0.009)
HighRank	-0.051*** (0.012)	-0.070*** (0.012)	-0.145*** (0.013)	-0.045*** (0.011)	-0.060*** (0.010)	-0.115*** (0.010)
LowRank*Trustee Fund	0.105*** (0.021)	0.132*** (0.022)	0.062*** (0.023)	0.080*** (0.017)	0.064*** (0.016)	0.060*** (0.016)
HighRank*Trustee Fund	-0.003 (0.018)	0.028* (0.018)	0.084*** (0.019)	0.018 (0.017)	0.047*** (0.019)	0.064*** (0.019)
Log(Plan Size)	-0.007*** (0.001)	-0.007*** (0.001)	-0.007*** (0.001)	-0.007*** (0.001)	-0.006*** (0.001)	-0.006*** (0.001)
Options	-0.001*** (0.000)	-0.001*** (0.000)	-0.001*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)	-0.002*** (0.000)
Fee	4.393*** (0.536)	3.938*** (0.539)	4.405*** (0.542)	4.762*** (0.538)	4.619*** (0.530)	4.757*** (0.529)
Turnover	0.008*** (0.001)	0.008*** (0.001)	0.009*** (0.001)	0.008*** (0.001)	0.007*** (0.001)	0.007*** (0.001)
Log(Fund Size)	-0.021*** (0.001)	-0.019*** (0.001)	-0.017*** (0.001)	-0.021*** (0.001)	-0.020*** (0.001)	-0.018*** (0.001)
Fund Age	0.000*** (0.000)	0.000*** (0.000)	0.000 (0.000)	0.000*** (0.000)	0.000*** (0.000)	0.000 (0.000)
Std. Dev.	-0.033 (0.120)	0.242** (0.120)	0.328*** (0.115)	-0.092 (0.131)	0.129 (0.130)	0.356*** (0.121)
Observations	99,967	99,967	99,967	100,299	100,299	100,299
Adj. R-Squared	0.080	0.088	0.086	0.077	0.084	0.086

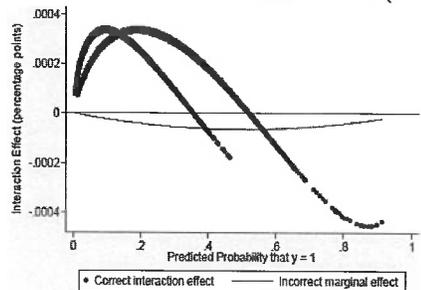
Panel A: Interaction Coefficients (LOW)



Panel B: Z-Statistics (LOW)



Panel C: Interaction Coefficients (HIGH)



Panel C: Z-Statistics (HIGH)

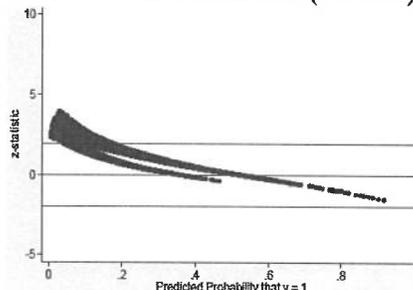


Figure 1: Interaction Effects by Ai-Norton

Figure A1: Robustness of Interaction Effect Between Trustee Indicator Variable and Performance Rank.

The following graphs display the interaction effects and corresponding z-statistics on the interaction variable between the trustee dummy and the performance ranks in Table A1 of the Appendix, estimated using Norton, Wang, and Ai (2004). The interaction effect is defined as the change in the predicted probability of deletion for a change in both the fund performance and the fund affiliation.

85

MORAL EQUILIBRIUM: STOCK BROKERS AND THE LIMITS OF DISCLOSURE

ROBERT A. PRENTICE*

At this writing, the SEC is pursuing one of the charges given it by Congress in the Dodd-Frank Act by attempting to decide whether to impose a uniform fiduciary obligation upon stock brokers and other professionals giving investment advice to customers or, among other options, to require brokers merely to disclose to customers that the brokers owe them no fiduciary duty. This Article draws upon the general literature of behavioral psychology and some fascinating new developments in the field of behavioral ethics to argue that disclosure in this narrow setting, and disclosure more generally as a default remedy for all manner of ills present in the financial system, are probably inadequate.

Introduction	1060
I. Fiduciary Duty or Mandatory Disclosure?	1061
II. Customer's Cognitive Limits and Behavioral Limitations....	1068
A. Bounded Rationality and Rational Ignorance	1069
B. Overoptimism and Overconfidence.....	1071
C. False Consensus Effect.....	1074
D. Insensitivity to Source of Information	1075
E. Probabilities and Future Events.....	1076
F. The Impact of Oral Communications	1077
G. Social Proof.....	1079
H. Difficulty of Judgment Correction.....	1080
1. Motivation to Make Correct Judgments	1080
2. Awareness of Potential Distorting Influences	1082
3. Awareness of the Direction and Magnitude of the Distorting Influences	1082
III. Bounded Ethicality, Moral License, and the Limitations of Disclosure	1085
A. Bounded Ethicality	1085
B. Moral Equilibrium	1094
1. Disclosure Distortion.....	1099
2. Disclosure, the Panhandler Effect, and Insinuation Anxiety	1103
Conclusion.....	1105

* Ed & Molly Smith Professor of Business Law, McCombs School of Business, University of Texas at Austin.

INTRODUCTION

The human mind is a wondrous thing. With fMRI technology and other recent breakthroughs, discoveries about how humans think are coming so quickly and furiously that in many fields such as psychology it is only slightly overly simplistic to say: "It's the brain, stupid." These developments are, and should be, rapidly infiltrating the literature of the law and of business ethics.¹

An important and ongoing public policy debate that should be substantially shaped by these new developments concerns whether the Securities and Exchange Commission (SEC) should impose a fiduciary duty upon stock brokers or, among other alternatives, require only that stock brokers disclose to their customers that they owe them no fiduciary duty.² While this paper does not take a position on many of the issues raised by this broad-ranging debate, it does draw substantially upon the literature of behavioral psychology, cognitive science, behavioral economics, behavioral ethics, and related fields³ to

1. See John Hasnas, Robert Prentice & Alan Strudler, *New Directions in Legal Scholarship: Implications for Business Ethics Research, Theory, and Practice*, 20 BUS. ETHICS Q. 503, 504 (2010) (exploring the implications of behavioral psychology research and related fields for legal and business ethics research).

2. The disclosure alternative has support beyond stock brokers themselves. See, e.g., Michael R. Newton, *Fiduciary Duties for Broker-Dealers* 15-16 (Apr. 25, 2011) (unpublished working paper), available at <http://ssrn.com/abstract=1822670>.

3. Over the years the author has frequently explored this literature and applied it to various policy issues, particularly in the securities law field. See, e.g., Robert A. Prentice, *The Case of the Irrational Auditor: A Behavioral Insight into Securities Fraud Litigation*, 95 NW. U. L. REV. 133, 218-19 (2000) [hereinafter Prentice, *Irrational Auditor*] (criticizing on behavioral grounds frequent court statements that auditors would never audit recklessly because it would be against their rational self-interest to do so); Robert A. Prentice, *Chicago Man, K-T Man, and the Future of Behavioral Law and Economics*, 56 VAND. L. REV. 1663, 1670 (2003) [hereinafter Prentice, *Chicago Man*] (defending the field of behavioral law and economics from critical attack); Robert Prentice, *Contract-Based Defenses in Securities Fraud Litigation: A Behavioral Analysis*, 2003 U. ILL. L. REV. 337, 419 [hereinafter Prentice, *Contract-Based Defenses*] (arguing that behavioral heuristics and biases can lead investors to unwisely forfeit legal protections when presented with form contracts containing liability disclaimers); Robert Prentice, *Enron: A Brief Behavioral Autopsy*, 40 AM. BUS. L.J. 417, 417-21 (2003) (exploring how various behavioral heuristics and biases might have contributed to the Enron debacle); Robert A. Prentice, *Ethical Decision Making: More Needed than Good Intentions*, 63 FIN. ANALYSTS J., no. 6, 2007 at 17, 17-18 (explaining how cognitive traps and behavioral heuristics can lead even well-intentioned financial analysts to make ethical missteps); Robert A. Prentice, *The SEC and MDP: Implications of the Self-Serving Bias for Independent Auditing*, 61 OHIO ST. L.J. 1597, 1602-04 (2000) [hereinafter Prentice, *SEC and MDP*] (exploring the implications of the self-serving bias for the policy arguments underlying audit regulation); Robert A. Prentice & Mark E. Roszkowski, "Tort Reform" and the Liability "Revolution": *Defending Strict Liability in Tort for Defective Products*, 27

argue that merely requiring an “NFD” (no fiduciary duty) disclosure is insufficient to adequately protect investors from the manifold conflicts of interest that are inherent in a stock broker-customer relationship.

Part I of this Article introduces the controversy. Part II explains why NFD disclosures are insufficient to enable customers to adequately protect themselves from stock broker abuses. Part III examines the issue from the stock brokers’ side of things and explains why NFD disclosures are insufficient to rein in stock broker misbehavior and, as it happens, gives additional reasons to suspect that NFD disclosures make it harder, not easier, for customers to protect themselves.

I. FIDUCIARY DUTY OR MANDATORY DISCLOSURE?

The subprime meltdown and credit crunch of 2007 and 2008 which led to “the Great Recession”⁴ had many causes and brought to light many abuses in the financial sector.⁵ As is often the case with such debacles, substantial financial regulatory overhaul followed in the form of the Dodd-Frank Act.⁶

An issue on the periphery of these scandals, but one of intrinsic importance, relates to the legal obligations of stock brokers to their customers. Stock brokers owe many duties to their customers, including a duty of best execution,⁷ a duty not to churn accounts,⁸ and a duty to

GONZ. L. REV. 251, 253–54 (1991–92) (arguing that limitations in human decision making indicate that more responsibility should be placed on companies that design consumer products and less on consumers who use them); Robert Prentice, *Whither Securities Regulation? Some Behavioral Observations Regarding Proposals for Its Future*, 51 DUKE L.J. 1397, 1399–1400 (2002) (criticizing on behavioral grounds a proposal to totally deregulate the securities business).

4. The Great Recession of 2008 to 2009 actually started in 2007. Chris Isidore, *It’s Official: Recession Since Dec. ‘07*, CNNMONEY.COM (Dec. 1, 2008, 5:40 PM), <http://money.cnn.com/2008/12/01/news/economy/recession/index.htm>. At this writing (August 2011), it looks like the Great Recession might be just the first of two “dips.” Graham Bowley, *Stocks Resume Free Fall on Fear over Economic and Credit Woes*, N.Y. TIMES, Aug. 9, 2011, at A1 (noting that a recent market sell-off was “creating the growing sense that the economy may be nearing a double-dip recession”).

5. There have been several accessible, popular accounts of the meltdown. See, e.g., CHARLES R. MORRIS, *THE TRILLION DOLLAR MELTDOWN* (2008); PAUL MUOLO & MATHEW PADILLA, *CHAIN OF BLAME: HOW WALL STREET CAUSED THE MORTGAGE AND CREDIT CRISIS* (2008); ROBERT POZEN, *TOO BIG TO SAVE?: HOW TO FIX THE U.S. FINANCIAL SYSTEM* (2010); ANDREW ROSS SORKIN, *TOO BIG TO FAIL* (2009).

6. Dodd-Frank Wall Street Reform and Consumer Protection Act (Dodd-Frank), Pub. L. No. 111-203, 124 Stat. 1376 (2010).

7. *Newton v. Merrill, Lynch, Pierce, Fenner & Smith, Inc.*, 135 F.3d 266, 270 (3d Cir. 1998).

8. *Nesbit v. McNeil*, 896 F.2d 380, 382 (9th Cir. 1990).

recommend only stocks that are “suitable” for the customers’ needs.⁹ What stock brokers do not owe their customers is a full-blown fiduciary obligation.¹⁰ This means that in a range of investment opportunities that would be suitable for a customer, the broker may recommend that opportunity which is the most lucrative to the broker even if it is the least beneficial to the customer.¹¹

Because investment advisors, who often perform similar services for customers, do owe their clients a fiduciary duty,¹² they have long been galled by the differential treatment of brokers. During the debate running up to passage of Dodd-Frank, the Obama administration proposed legislation to impose a duty “to act solely in the interest of the customer or client without regard to the financial or other interest of the broker, dealer or investment advisor providing the advice,”¹³ thus

9. FIN. INDUS. REGULATORY AUTH., FINRA MANUAL, NASD Rule 2310 (1996) [hereinafter FINRA MANUAL]. See generally Frederick Mark Gedicks, *Suitability Claims and Purchases of Unrecommended Securities: An Agency Theory of Broker-Dealer Liability*, 37 ARIZ. ST. L.J. 535 (2005) (discussing the suitability concept generally).

10. Ronald J. Colombo, *Trust and the Reform of Securities Regulation*, 35 DEL. J. CORP. L. 829, 863 (2010); Arthur B. Laby, *The Fiduciary Obligation as the Adoption of Ends*, 56 BUFF. L. REV. 99, 161–62 (2008). FINRA rules generally require stock brokers to act reasonably, but not necessarily in the best interests of their customers. See FINRA MANUAL, *supra* note 9.

11. See Steven D. Irwin et al., *Wasn't My Broker Always Looking Out for My Best Interests? The Road to Become a Fiduciary*, 12 DUQ. BUS. L.J. 41 (2009) (containing a lengthy comparison of the fiduciary duty versus suitability duty approaches); see also Newton, *supra* note 2, at 2–7 (comparing differences in standards and regulation between investment advisors and stock brokers).

12. Broker-dealers and investment advisors both provide retail investors with investment advice in order for the investors to achieve their financial goals. Despite the similar roles of providing the benefit of investment advice for retail investors, broker-dealers are regulated under the Securities Exchange Act of 1934, 15 U.S.C. §§ 78a–78nn (2006), while investment advisors are regulated under the Investment Advisers Act of 1940, 15 U.S.C. §§ 80b-1 to -21 (2006). In fact, Congress intentionally excluded broker-dealers from regulation under the Investment Advisers Act. *Id.* § 80b-2(a)(11)(C). Although the roles of investment advisor (giving investment advice for compensation) and stock broker (executing transactions on behalf of a client with or without providing investment advice) may have been more clearly differentiated in the 1930s, in recent years the practices have tended to converge. Brokers often provide investment advice to their customers and today some professionals are dual-registrants: investment advisors who are also registered as broker-dealers and vice versa. SEC. & EXCH. COMM'N, STUDY ON INVESTMENT ADVISERS AND BROKER-DEALERS 12–13 (2011).

13. Press Release, U.S. Dep't of the Treasury, Additional Improvements to Financial Markets Regulation (July 10, 2009), <http://www.treasury.gov/press-center/press-releases/Documents/tg205071009.pdf>. This proposal was incorporated into the original House version of Dodd-Frank in § 7103. See The Wall Street Reform and Consumer Protection Act of 2009, H.R. 4173, 111th Cong. § 7103(a)(3) (2009).

raising the bar for stock brokers and leveling the playing field between the two professions.

As enacted, Dodd-Frank did not impose a uniform fiduciary duty upon all professionals giving financial advice. Rather, it merely required the SEC to devise a plan for investors to receive clear disclosures from their investment advice professionals concerning any material conflicts of interest.¹⁴ Dodd-Frank charged the SEC with studying the issues and then deciding whether to resort to the trusted default of disclosure, regulating through other methods such as imposing trading limitations, or imposing a uniform fiduciary duty upon investment advice professionals, including broker-dealers.

In studying the issues, the SEC has run into many thorny questions. On April 7, 2011, the SEC indicated that it would not undertake rule-making regarding a potential uniform fiduciary standard until later in 2011.¹⁵ So, it remains quite possible that disclosure may turn out to be the preferred remedy. Having to make NFD disclosures would no doubt be awkward for stock brokers, but they generally prefer a disclosure requirement to being saddled with a full-blown fiduciary duty.

Reliance upon mandatory disclosure would certainly be a business-as-usual approach. The sunshine of disclosure is often touted as the best “disinfectant,”¹⁶ and there is little doubt that disclosure has been a popular and generally successful remedy in federal securities law from the very beginning in 1933.¹⁷ Mandatory disclosure by issuers in 10-Ks,

14. See Dodd-Frank, Pub. L. No. 111-203, § 913(g), 124 Stat. 1376, 1828 (requiring the SEC to “facilitate the provision of simple and clear disclosures to investors regarding the terms of their relationships with brokers, dealers, and investment advisers, including any material conflicts of interest”).

15. *UPDATE 1-US SEC Sees Fiduciary Standard Rules Late in 2011*, REUTERS (Apr. 7, 2011), <http://www.reuters.com/assets/print?aid=USN0720336220110407>.

16. LOUIS D. BRANDEIS, *OTHER PEOPLE’S MONEY AND HOW THE BANKERS USE IT* 92 (2d prtg. 1914) (“Sunlight is . . . the best of disinfectants; electric light the most efficient policeman.”).

17. Disclosure has been the most popular remedy in federal securities laws, beginning with the 1933 Securities Act and the 1934 Securities Exchange Act. See 1 LOUIS LOSS & JOEL SELIGMAN, *SECURITIES REGULATION* 29 (3d ed. 1998) (noting that the “recurrent theme” throughout federal securities law has been “disclosure, again disclosure, and still more disclosure”). In enacting Sarbanes-Oxley, Congress relied significantly upon disclosure as well. See Susanna Kim Ripken, *The Dangers and Drawbacks of the Disclosure Antidote: Toward a More Substantive Approach to Securities Regulation*, 58 BAYLOR L. REV. 139, 143–44 (2006) (noting the tendency of Sarbanes-Oxley to use disclosure rather than more substantive solutions).

10-Qs, 8-Ks, registrations statements, and the rest has been particularly beneficial.¹⁸

Given the benefits of mandatory disclosure by issuers, it is not surprising that disclosure has become the default remedy in many areas of securities regulation, as well as many other areas of public policy. Still, there are grounds for reservation. Antibiotics are wonderful things, but as evolving staph infections have recently demonstrated, their efficacy has limits.¹⁹ Although there is no doubt that the broad corporate disclosure requirements of the 1933 and 1934 Acts are beneficial in a number of ways, automatic resort to disclosure as a

18. There is substantial empirical evidence that a rigorous securities law regime is associated with cheaper capital, greater development of stock markets, and more vigorous economic growth, among other benefits. See FRANK B. CROSS & ROBERT A. PRENTICE, LAW AND CORPORATE FINANCE 152-89 (2007) (surveying the empirical literature and adding to it). The authors found rough evidence that rigorous securities laws "may . . . produce significant economic benefits, measured in the trillions of dollars." *Id.* at 188. In addition to strict antifraud rules and prohibitions against insider trading, mandatory issuer disclosure can be an important feature of such a rigorous regime. See Frank B. Cross & Robert A. Prentice, *The Economic Value of Securities Regulation*, 28 CARDOZO L. REV. 333, 365-89 (2006) (surveying empirical literature, supplementing it, and concluding that "[t]here is an enormous weight of rigorous study of the effects of securities regulation generally and specific aspects such as mandatory disclosure rules or insider trading prohibitions" and that those studies provide "strong empirical evidence of the general beneficence of comprehensive securities regulation"). In a recent study of a natural experiment arising from EU nations' adoption of new market abuse and *transparency regulation* directives, Christensen and colleagues found that market liquidity increased and firms' cost of capital decreased as countries implemented and enforced the directives. Hans B. Christensen et al., *Capital-Market Effects of Securities Regulation: The Role of Prior Regulation, Implementation and Enforcement* 3-5 (Oct. 2011) (unpublished working paper), available at <http://ssrn.com/abstract=1745105>. Empirical evidence indicates that issuers in regimes of mandatory disclosure backed by meaningful punishments for fraud are able to raise funds at a cheaper cost than in other regimes because investors can be more confident in making their decisions to put money into a firm. See Allen Ferrell, *The Case for Mandatory Disclosure in Securities Regulation around the World*, 2 BROOK. J. CORP., FIN. & COM. L. 81, 93-96 (2007). Certainly investors can invest and creditors can lend with much greater confidence if they know that strict antifraud rules apply and will be used to punish those who would deceive them. See Stephen J. Choi, *Selective Disclosures in the Public Capital Markets*, 35 U.C. DAVIS L. REV. 533, 551 (2002). Otherwise, they will have to discount the value of *all* firms. *Id.* The standardization and comparability introduced by mandatory issuer disclosures increase the value of all corporate information. Gerard Hertig et al., *Issuers and Investor Protection*, in THE ANATOMY OF CORPORATE LAW: A COMPARATIVE AND FUNCTIONAL APPROACH 193, 206 (2004). All investors will find it more cost effective to have issuers disclose information than to search it out themselves. See Roberta Romano, *Empowering Investors: A Market Approach to Securities Regulation*, 107 YALE L.J. 2359, 2417 (1998).

19. See BRAD SPELLBERG, RISING PLAGUE: THE GLOBAL THREAT FROM DEADLY BACTERIA AND OUR DWINDLING ARSENAL TO FIGHT THEM 43-44 (2009).

remedy for every problem is inadequate²⁰ and occasionally counterproductive.²¹

In the financial and securities worlds, overreliance on disclosure arose as the efficient markets hypothesis and the rational man dogma of modern financial economics came to the fore.²² If, as economists have tended to assume, the markets are made up of fully rational investors who make optimal resource allocation and wealth maximizing decisions, then full disclosure is an appropriate remedy for all that ails the markets.²³ Indeed, it might be the only remedy needed.

Unfortunately, research in the fields of behavioral decision theory, cognitive psychology, behavioral economics, behavioral finance, brain science, and others have largely decimated any notion that people, particularly when they act as investors, are the embodiment of the “Chicago Man” model of rationality.²⁴ Although there is still some

20. The limits of disclosure have been written about in several contexts lately. See, e.g., Geoffrey A. Manne, *The Hydraulic Theory of Disclosure Regulation and Other Costs of Disclosure*, 58 ALA. L. REV. 473 (2007) (discussing many problems with mandatory disclosure generally, and as related to executive compensation specifically, especially including unintended consequences caused by unforeseen reactions of those regulated); Jennifer S. Martin, *The House of Mouse and Beyond: Assessing the SEC's Efforts to Regulate Executive Compensation*, 32 DEL. J. CORP. L. 481, 513–14 (2007) (arguing that merely disclosing executive pay to shareholders would be inadequate to solve the problem of excessive compensation because (a) managers influence the compensation process; (b) shareholders upset by the disclosures have little recourse beyond selling their shares; (c) shareholders have few means to make a board accountable for its compensation decisions; and (d) disclosure is an inadequate substitute for substantive regulation); Ripken, *supra* note 17, at 146–49 (arguing that substantive regulation would often be a better policy prescription than more mandated disclosure).

21. Regarding the problem of excessive executive compensation, for example, it has been observed that every time disclosure increases, so does executive pay. See Floyd Norris, *Which Bosses Really Care If Shares Rise?*, N.Y. TIMES, June 2, 2006, at C1. This is likely true, in part, because people like to compare salaries and with more disclosure, executives are more likely to find out that there are others out there making more than they are. *Id.*; see also ARCHON FUNG ET AL., FULL DISCLOSURE: THE PERILS AND PROMISE OF TRANSPARENCY 7 (2007) (noting that “targeted transparency policies can also do more harm than good” if information is obsolete, confusing, inaccurate or distorted).

22. Emiliios Avgouleas, *What Future for Disclosure as a Regulatory Technique? Lessons from the Global Financial Crisis and Beyond* 2 (Mar. 26, 2009) (unpublished working paper), available at <http://ssrn.com/abstract=1369004>.

23. See *id.* at 6–8.

24. FUNG ET AL., *supra* note 21, at 16 (“[C]ognitive psychologists and economists have persuasively challenged the notion that individuals and groups automatically use available information to make rational choices, documenting an array of cognitive shortcuts that distort the processing of new data.”); Russell Korobkin, *What Comes after Victory for Behavioral Law and Economics?* 3–4 (UCLA School of Law, Law-Econ Research Paper No. 11-10, 2011), available at <http://ssrn.com/abstract=1787070> (“I am declaring victory in the battle for the

essential truth in the notion that man is a rational animal and still some usefulness to this “Chicago Man” model,²⁵ the evidence regarding the limitations of human decision making, some of which will be summarized later in this Article, is overwhelming,²⁶ meaning that the pins have been knocked out from under the rationale for using disclosure as the main tool of financial and securities regulation.

Because of these developments in the science of human information processing and decision making, there is some reason to doubt whether an NFD disclosure would be a remedy sufficient to protect the interests of customers of stock brokers. One of the reasons that Congress ordered the SEC to study the issue was concern that retail investors were not aware of the differences between the duties owed by investment advisors and those owed by stock brokers and that they were also blissfully ignorant of the implications of those differences. Indeed, a recent survey found that sixty-six percent of investors thought that their brokers owed them a fiduciary duty.²⁷

How to reform the status quo in order to adequately protect investors is a policy question of significant import. Conflicts of interest abound in the securities industry. Stock brokers are certainly no exception, for they labor under numerous conflicts of interest.²⁸ Among the many ways that brokers can profit at the expense of the best

methodological soul of the law and economics discipline.”). *See generally* Prentice, *Chicago Man*, *supra* note 3 (contrasting the evidence supporting the competing behavioral and economic models of human behavior).

25. *See* Daniel McFadden, *Rationality for Economists?*, 19 J. RISK & UNCERTAINTY 73, 76, 97 (1999). McFadden labels the rational man of traditional economic models “Chicago man” and argues that “[c]onfronted with the accumulated experimental evidence, economists must recognize that the Chicago-man model does not apply universally, or even regularly, to choices made in non-market contexts.” *Id.*

26. *See, e.g.*, Michael P. Vandenbergh et al., *Regulation in the Behavioral Era*, 95 MINN. L. REV. 715, 716 (2011) (“The fact that individuals deviate in predictable ways from neoclassical assumptions of rationality has been widely recognized in the academic literature and has become well known to the public.”). *See generally* MAX H. BAZERMAN & DON A. MOORE, *JUDGMENT IN MANAGERIAL DECISION MAKING* (7th ed. 2009); REID HASTIE & ROBYN M. DAWES, *RATIONAL CHOICE IN AN UNCERTAIN WORLD* (2001); SCOTT PLOUS, *THE PSYCHOLOGY OF JUDGMENT AND DECISION MAKING* (1993).

27. Presentation, Infogroup/ORC, U.S. Investors & The Fiduciary Standard: A National Opinion Survey (Sept. 15, 2010) [hereinafter Infogroup/ORC], *available at* http://www.cfp.net/downloads/US_Investors_Opinion_Survey_2010-09-16.pdf. Also, thirty-four percent did not understand the differing roles of brokers and investment advisors. *Id.*

28. *See* Christopher Tarver Robertson, *Biased Advice*, 60 EMORY L.J. 653, 674–75 (2011) (noting the conflicts of interest stock brokers face, including a strong bias toward frequent trades to generate fees).

interests of their customers²⁹ are churning,³⁰ reverse churning,³¹ excessive mark-ups and commissions,³² failing best execution,³³ failing to disclose market-maker status,³⁴ price manipulation,³⁵ unauthorized trading,³⁶ selling unsuitable securities,³⁷ and operating boiler rooms.³⁸

It is true that because of these and other unsavory practices that can arise out of the heavily conflicted broker-customer relationship, there are a raft of industry standards and state and federal regulations that govern the brokerage profession. Federal law backs up the Financial Institution Regulatory Authority's (FINRA's) Rule 2310,

29. See James J. Angel & Douglas McCabe, Ethical Standards for Stockbrokers: Fiduciary or Suitability? 5-9 (Sept. 30, 2010) (unpublished working paper), available at <http://ssrn.com/abstract=1686756> (discussing the many ways that brokers may profit from their clients and the many conflicts of interest inherent in the relationship).

30. Churning, of course, is overtrading a customer's account for the primary purpose of generating commissions. See, e.g., *Kravitz v. Pressman, Frohlich & Frost, Inc.*, 447 F. Supp. 203, 210-11 (D. Mass. 1978) (finding a classic case of churning). See generally Craig McCann, *Churning*, J. LEGAL ECON., Spring/Summer 1999, at 49 (analyzing churning through an economic lens).

31. Reverse churning occurs when compensation is based upon the amount of assets in the account and trades are uncompensated, which often leads to undertrading or an account being ignored altogether. See Arthur B. Laby, *Fiduciary Obligations of Broker-Dealers and Investment Advisers*, 55 VILL. L. REV. 701, 740 (2010) (discussing reverse churning).

32. See, e.g., *Baber v. First Republic Grp., L.L.C.*, No. C 06-3076-MWB, 2008 U.S. Dist. LEXIS 45180 (N.D. Iowa June 6, 2008) (claim for excessive mark-ups survived a motion to dismiss).

33. Brokers executing trades on their customers' behalfs have an obligation to execute trades at the best net price considering all relevant circumstances. See, e.g., *Newton v. Merrill, Lynch, Pierce, Fenner & Smith, Inc.*, 135 F.3d 266, 270 (3d Cir. 1998) (claim that defendant-brokers failed their best execution responsibilities survived motion to dismiss).

34. Brokers' interests can conflict with customers' interests in a number of ways, and customers must know if their brokers are acting as agents, principals, or market makers to have a chance to gauge those conflicts. See, e.g., *Chasins v. Smith, Barney & Co.*, 438 F.2d 1167 (2d Cir. 1970) (firm punished for failing to disclose market-maker status).

35. See, e.g., *Pagel, Inc. v. SEC*, 803 F.2d 942 (8th Cir. 1986) (finding price manipulation by a stock broker).

36. Carrying out unauthorized trades on nondiscretionary accounts is sometimes called writing "wooden tickets." *In re Sterling Foster & Co.*, 222 F. Supp. 2d 216, 232 (E.D.N.Y. 2002).

37. See, e.g., *Erdos v. SEC*, 742 F.2d 507 (9th Cir. 1984) (finding against defendant-broker in an unsuitability case).

38. See, e.g., *SEC v. Rabinovich & Assocs. LP*, No. 07 CIV. 10547(GEL), 2008 U.S. Dist. LEXIS 93595 (S.D.N.Y. Nov. 18, 2008) (finding that defendants ran a classic boiler-room operation).

which imposes upon brokers a suitability responsibility.³⁹ As noted earlier, in recommending purchases, sales, or exchanges to their customers, brokers are to have reasonable grounds to believe that a stock is suitable for the customer.⁴⁰ However, as also noted earlier, if there are multiple options that are all suitable for customers, brokers may choose the stock that creates the greatest compensation for them, even if it is the poorest suitable choice for the customer. While this would violate a fiduciary duty if the brokers owed one to the customers, since no such duty is currently owed, it is perfectly fine under the existing regime. Would a required NFD disclosure sufficiently protect customers and rein in brokers?

II. CUSTOMER'S COGNITIVE LIMITS AND BEHAVIORAL LIMITATIONS

If people had full information and were fully rational in how they processed it, disclosure might be the only remedy needed in numerous situations. Such is not the case, however, as the discussion in this Part illustrates. Some problems in relying solely upon disclosure to enable investors to protect themselves are primarily cognitive in nature, stemming from the confusing form and complex nature of disclosures.⁴¹ Information overload is often a significant problem as well.⁴²

39. See generally 7 LOUIS LOSS & JOEL SELIGMAN, *SECURITIES REGULATION* 3885-93 (3d ed. 2004).

40. FINRA MANUAL, *supra* note 9.

41. Disclosure is often in complicated legalese rather than plain English, making it difficult for even sophisticated investors to understand it well enough to protect themselves. See, e.g., Alan B. Levenson, *The Role of the SEC as a Consumer Protection Agency*, 27 BUS. LAW. 61, 68 (1971) (noting problems caused by disclosures in the form of boilerplate legalese); Ripken, *supra* note 17, at 186 (noting difficulties with disclosure documents couched in legalese). This has been a particular problem in the securities business, where issuer disclosure has been described as "static" and "obtuse." Susan M. Brunka, *XBRL and the SEC: How the Commission Uses Interactive Data to Investigate Illegal Stock Options Backdating and What Interactive Data Means for the Future of Federal Securities Law Enforcement*, 44 CAL. W. L. REV. 87, 88 (2007). The SEC has sensibly attacked this problem with a range of "plain English" initiatives. See, e.g., Robert Sprague & Corey Ciochetti, *Preserving Identities: Protecting Personal Identifying Information through Enhanced Privacy Policies and Laws*, 19 ALB. L.J. SCI & TECH. 91, 132 (2009) (discussing some SEC "plain English" initiatives); Michael G. Byers, Note, *Eschew Obfuscation—The Merits of the SEC's Plain English Doctrine*, 31 U. MEM. L. REV. 135, 152-56, 165-68 (2000) (evaluating SEC initiatives). However, even plain English does not help when the matters being disclosed are overly complex, as is often the case in the securities field. See, e.g., Alison Grey Anderson, *The Disclosure Process in Federal Securities Regulation: A Brief Review*, 25 HAST. L.J. 311, 325 (1973-74) (noting that prospectus disclosure is so complex that it is difficult for investors to detect even blatant fraud); Martin, *supra* note 20, at 525 ("[T]he complexity of modern corporate enterprises

The focus of this discussion, however, will be primarily upon insights derived from behavioral psychology and closely related fields. I have elsewhere labored mightily to demonstrate that cognitive and behavioral limitations often cause stock brokers' customers to sign away beneficial legal protections.⁴³ Many of those same limitations will also prevent customers given an NFD disclosure from adequately protecting themselves from the self-serving actions of brokers.

A. Bounded Rationality and Rational Ignorance

Many who tout disclosure as the only required remedy for a variety of ills assume that investors are *homo economicus*, fully rational investors who gather and utilize all information that is relevant to an investment decision. That assumption is almost always at least somewhat inaccurate. As Herbert Simon noted many years ago, because they seldom have complete and perfectly accurate information and never have perfect capacity to process that information, people are "*intendedly* rational, but only *limitedly* so."⁴⁴ Therefore, the fact that investors are informed that their brokers owe them no fiduciary duty

makes clear that accurate disclosure that is understandable to shareholders is a daunting task for corporate officers.").

42. Information overload occurs when so much information is disclosed that it is less helpful to investors than much less information. See generally Troy A. Paredes, *Blinded by the Light: Information Overload and Its Consequences for Securities Regulation*, 81 WASH. U. L.Q. 417 (2003). In any disclosure system, information overload is a potential problem. Too much detailed information can swamp readers who are unable to discern the important information from the masses of data points they are provided. See generally Anderson, *supra* note 41, at 325 (discussing information overload in securities law more generally); Martin, *supra* note 20, at 525-26 (discussing information overload with specific regard to executive compensation disclosures). "When faced with too much data, people have a tendency to become distracted by less relevant information and to ignore information that may turn out to be highly relevant." Ripken, *supra* note 17, at 160. More information can actually lead to worse decisions. See Jacob Jacoby, *Is It Rational to Assume Consumer Rationality? Some Consumer Psychological Perspectives on Rational Choice Theory*, 6 ROGER WILLIAMS U. L. REV. 81, 133 (2000) ("[W]hen it comes to making decisions, more information is not necessarily better, and sometimes may be considerably worse."). This is a problem for sophisticated investors as well as lay investors. Steven L. Schwarcz, *Temporal Perspectives: Resolving the Conflict between Current and Future Investors*, 89 MINN. L. REV. 1044, 1069 (2005) ("Even the most sophisticated investors can suffer information overload."). Striking the proper balance is often a difficult task.

43. Prentice, *Contract-Based Defenses*, *supra* note 3, at 338, 358-59.

44. HERBERT A. SIMON, *ADMINISTRATIVE BEHAVIOR* xxiv (2d ed. 1957). See generally Amanda P. Reeves & Maurice E. Stucke, *Behavioral Antitrust*, 86 IND. L.J. 1527, 1532-35 (2011) (containing a recent discussion of the concept of bounded rationality).

does not mean that the investors will fully understand the disclosure or, even if they do, respond to it appropriately.

In addition to being boundedly rational, in most settings people remain rationally ignorant.⁴⁵ Because they do not have unlimited time to gather or to process information, they do not do so. Because they have numerous demands on their time and attention, they will generally “satisfice” rather than optimize their information gathering and decision making.⁴⁶ This tendency is so strong that “even professional security analysts deciding on which securities to select do not acquire most (or even much) of the information available.”⁴⁷

Because of rational ignorance and other factors, lay investors typically do not read disclosure documents when investing in securities,⁴⁸ meaning that investors who receive a document indicating

45. See Shawn J. Bayern, *Rational Ignorance, Rational Closed-Mindedness, and Modern Economic Formalism in Contract Law*, 97 CALIF. L. REV. 943, 943 (2009) (“Rational ignorance is the familiar notion that it would be foolish for an individual to gather and process all possible information.”).

46. ROBYN M. DAWES, RATIONAL CHOICE IN AN UNCERTAIN WORLD 50–56 (Jerome Kagen ed., 1988) (noting implications of the tendency of people to “satisfice” rather than to optimize); Herbert A. Simon, *Rational Choice and the Structure of the Environment*, 63 PSYCHOL. REV. 129 (1956) (noting that people tend to “satisfice” rather than optimize when making decisions).

47. Jacoby, *supra* note 42, at 132–33. See generally Jacob Jacoby et al., *Effectiveness of Security Analyst Information Accessing Strategies: A Computer Interactive Assessment*, 1 COMPUTERS IN HUM. BEHAV. 95 (1985). Studies show that “operating under conditions that have both financial and ego consequences and where information acquisition costs are virtually zero, even professional security analysts deciding on which securities to select do not acquire most (or even much) of the information available.” Jacoby, *supra* note 42, at 132–33. Naturally, lay investors gather even less information that is available to them. Professors Ben-Shahar and Schneider observe that people do not like making financial decisions and avoid gathering information that is necessary to make those decisions. Omri Ben-Shahar & Carl E. Schneider, *The Failure of Mandated Disclosure*, 159 U. PA. L. REV. 647, 728 (2011) For example, only forty-two percent of workers have even attempted to calculate how much money they will need to save for their retirement. Ruth Helman et al., *Encouraging Workers to Save: The 2005 Retirement Confidence Survey*, EBRI ISSUE BRIEF (Emp. Benefits Research Inst.), Apr. 2005, at 6. Financial information that is disclosed to investors is often beyond their ken. Even when regulators require disclosure to be *simplified*, it is very often beyond the ability of most investors to meaningfully comprehend. See generally John Beshears et al., *How Does Simplified Disclosure Affect Individuals’ Mutual Fund Choices?* (Harv. Univ. Kennedy Sch. of Gov’t., Working Paper No. RWP09 16, 2009), available at <http://ssrn.com/abstract=1400943> (giving mutual-fund disclosures as an example). Even if they understand it in the first place, they are unlikely to remember it correctly. “Studies of informed consent . . . find that people correctly remember about one-third of the basic information” that is conveyed to them. Ben-Shahar & Schneider, *supra*, at 719.

48. See, e.g., Homer Kripke, *The Myth of the Informed Layman*, 28 BUS. LAW. 631 (1973) (arguing that mandatory securities law disclosure is not justified

that their stock broker owes them no fiduciary duty often (a) will not even read it and therefore go on assuming that they are, in fact, owed a fiduciary duty by their brokers, as most investors currently believe,⁴⁹ or (b) if they do read it, they will not go to the trouble of figuring out what the heck it means not to be owed a fiduciary duty.⁵⁰ And this is a rational, if not optimal, course of action for them. After all, they have lives to live and cannot spend all of their time gathering information about just one of hundreds of decisions they must make during the course of a day.

Furthermore, if *no* stock broker owes them a fiduciary duty, what are they to do with the information anyway? The disclosure that this particular broker does not owe them a fiduciary duty does not assist them in choosing among potential brokers if none of them owes a fiduciary obligation.

B. Overoptimism and Overconfidence

Even if investors read NFD warnings and understand roughly what they mean, they likely will not react in a fully rational manner. Humans lean toward irrational optimism in many settings, tending to believe that the bad things that happen to other people (divorce, cancer, crooked stock brokers) will not happen to them.⁵¹ Manufacturers have often

because most investors do not read most documents filed with the SEC); Baruch Lev & Meiring de Villiers, *Stock Price Crashes and 10b-5 Damages: A Legal, Economic, and Policy Analysis*, 47 STAN. L. REV. 7, 19 (1994) (“[M]ost investors do not read, let alone thoroughly analyze, financial statements, prospectuses, or other corporate disclosures.”).

49. See *supra* note 27 and accompanying text. It is not surprising that most investors think that stock brokers owe them a fiduciary duty: “a sizable percentage of mortgage borrowers believe that their lenders are legally obliged to offer them the best possible rate.” James Surowiecki, *The Warren Court*, NEW YORKER, June 13 & 20, 2011, at 48.

50. Owen D. Jones, *Time-Shifted Rationality and the Law of Law’s Leverage: Behavioral Economics Meets Behavioral Biology*, 95 NW. U. L. REV. 1141, 1151 (2001) (explaining that given the high costs of gathering *all* relevant information, time constraints on studying that information, and other human limitations, “it [is] clear that in some cases, paradoxically, *it would be irrational to become fully informed*” (emphasis added)). “Typical consumers do not ask for or read their contracts pre-purchase.” Amy J. Schmitz, *Pizza-box Contracts: True Tales of Consumer Contracting Culture*, 45 WAKE FOREST L. REV. 863, 863 (2010).

51. See generally Lynn A. Baker & Robert E. Emery, *When Every Relationship Is above Average: Perceptions and Expectations of Divorce at the Time of Marriage*, 17 LAW & HUM. BEHAV. 439, 443 (1993) (finding that most newlyweds know that half of married couples will divorce but place their own chance at zero); Jon D. Hanson & Douglas A. Kysar, *Taking Behavioralism Seriously: Some Evidence of Market Manipulation*, 112 HARV. L. REV. 1420, 1512 (1999) (“Considerable evidence suggests that smokers perceive smoking as significantly less risky for themselves than

taken advantage of this inclination in order to successfully sell risky products,⁵² and the same abuses exist in investment sales as well.⁵³ Because of their overoptimism, investors tend to believe that they will avoid being ripped off and will achieve good investment results, even as other investors might suffer mightily at the hands of crooked stock brokers.⁵⁴

Overoptimism tends to be reinforced by another common human tendency—that of overconfidence.⁵⁵ Studies indicate that people tend, in mathematically impossible percentages, to believe that they are above average in driving, auditing, and teaching.⁵⁶ This overconfidence leads many people to also believe that they are better than average investors,⁵⁷ even though it is quite clear that most amateur investors are

for other smokers, that smokers view their own risks as *not* significantly higher than those for non-smokers, and that smokers tend to underestimate the actual risks to themselves.”). This overoptimism is also referred to as the “illusion of hope.” Jeffrey B. Coopersmith, Comment, *Refocusing Liquidated Damages Law for Real Estate Contracts: Returning to the Historical Roots of the Penalty Doctrine*, 39 EMORY L.J. 267, 268 (1990).

52. Jon D. Hanson & Douglas A. Kysar, *Taking Behavioralism Seriously: The Problem of Market Manipulation*, 74 N.Y.U. L. REV. 630, 729 (1999).

53. See Werner F.M. De Bondt, *A Portrait of the Individual Investor*, 42 EUR. ECON. REV. 831, 839 (1998) (survey finding that investors tend toward overoptimism); see also HERSH SHEFRIN, BEYOND GREED AND FEAR: UNDERSTANDING BEHAVIORAL FINANCE AND THE PSYCHOLOGY OF INVESTING 18–20 (2000); Jennifer Arlen, Comment, *The Future of Behavioral Economic Analysis of Law*, 51 VAND. L. REV. 1765, 1773 (1998).

54. See Henry T. C. Hu, *Faith and Magic: Investor Beliefs and Government Neutrality*, 78 TEX. L. REV. 777, 861 (2000) (“Investors have a patently unrealistic view of the true downside [of investing in the stock market].”).

55. DAVID BROOKS, THE SOCIAL ANIMAL: THE HIDDEN SOURCES OF LOVE, CHARACTER AND ACHIEVEMENT 218 (2011) (“The human mind is an overconfidence machine.”); Christoph Merkle & Martin Weber, *True Overconfidence: The Inability of Rational Information Processing to Account for Apparent Overconfidence*, 116 ORG. BEHAV. & HUM. DECISION PROCESSES 262, 263, 271 (2011) (finding apparent overconfidence is really caused by true overconfidence, not by an inability to process information); Rick K. Wilson, *The Contribution of Behavioral Economics to Political Science*, 14 ANN. REV. POL. SCI. 201, 215 (2011) (“Overconfidence is a well-known problem.”).

56. See MAX BAZERMAN, JUDGMENT IN MANAGERIAL DECISION MAKING 95 (4th ed. 1998) (“[P]eople have been found to perceive themselves as being better than others across a number of traits, including honesty, cooperativeness, rationality, driving skill, health, and intelligence.”); SAM HARRIS, THE MORAL LANDSCAPE: HOW SCIENCE CAN DETERMINE HUMAN VALUES 188 (2010) (“[T]here is almost nothing more common than the belief that one is above average in intelligence, wisdom, honesty, etc.”).

57. See SHEFRIN, *supra* note 53, at 132–33 (noting studies showing investor overconfidence); De Bondt, *supra* note 53, at 831, 834–36 (finding overconfidence in investors generally, as well as overoptimism, a tendency to discount diversification, and a rejection of the notion that there is a positive tradeoff between risk and return).

poor traders.⁵⁸ Even investors who are experienced and relatively successful tend toward overconfidence.⁵⁹

In this setting, many investors will be optimistic that the frauds and unfair treatment of investors that they read about in the media will happen to others, not to them. This is especially so because investors tend to be confident of their ability to hire competent and honest brokers. These tendencies toward overoptimism and overconfidence affect risk perception, causing many people to underestimate the extent to which they are at risk of, for example, being ripped off,⁶⁰ especially because people tend to overestimate their ability to evaluate the trustworthiness of others.⁶¹ The overarching confidence that so many investors had in Bernie Madoff is illustrative.

A reinforcing phenomenon, called the "illusion of control,"⁶² leads many people to feel that they can exert control over events that actually are not within their control, including purely random events, causing them to believe that their chances of avoiding injury, for example, from being mistreated by their brokers, are "inappropriately higher than the objective probability would warrant."⁶³ They will often not be sufficiently concerned about the NFD warning because people

58. Studies of overconfidence have become very common in business in areas such as finance and management. See Ulrike Malmendier & Geoffrey Tate, *Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction*, 89 J. FIN. ECON. 20, 42 (2008); Menachem Brenner et al., *Ambiguity and Overconfidence* 3 (Feb. 24, 2011) (unpublished working paper), available at <http://ssrn.com/abstract=1773663>.

59. GARY BELSKY & THOMAS GILOVICH, *WHY SMART PEOPLE MAKE BIG MONEY MISTAKES—AND HOW TO CORRECT THEM: LESSONS FROM THE NEW SCIENCE OF BEHAVIORAL ECONOMICS* 152 (1999) (noting that many studies over the years have shown that a wide variety of professionals, including securities analysts, tend to be overconfident).

60. See, e.g., Neil D. Weinstein, *Unrealistic Optimism about Future Life Events*, 39 J. PERSONALITY & SOC. PSYCHOL. 806 (1980).

61. See, e.g., Peter J. DePaulo et al., *Lying and Detecting Lies in Organizations*, in *IMPRESSION MANAGEMENT IN THE ORGANIZATION* 377, 386 (Robert A. Giacalone & Paul Rosenfeld eds., 1989).

62. People apparently have a psychological need to believe that they can control their surroundings, including random events. Studies show that people are willing to bet more in a lottery if they choose the numbers themselves than if someone else chooses the numbers, even though the statistical chance of winning is unaffected. People also tend to throw dice harder if they want high numbers and softer if they want low numbers. See generally BAZERMAN, *supra* note 56, at 95–96 (discussing the "illusion of control"); Ellen J. Langer, *The Illusion of Control*, 32 J. PERSONALITY & SOC. PSYCHOL. 311 (1975).

63. PETER ASCH, *CONSUMER SAFETY REGULATION* 76 (1988) (quoting Langer, *supra* note 62, at 313).

optimistically tend to believe that the terms of contracts they have signed are more favorable to them than they actually are.⁶⁴

C. False Consensus Effect

The false consensus effect is the tendency of people to believe that others see the world as they do.⁶⁵ One ramification of the false consensus effect is that investors who are honest and would never consider misleading or taking advantage of another may naively believe the same of their stock brokers, who may not share the same values.⁶⁶ In other words, honest people do not tend to expect to be treated dishonestly. In her book on the Bernie Madoff scandal, Diana Henriques quoted Nobel laureate Elie Wiesel, one of Madoff's victims: "Wiesel offered his own explanation for how the scandal could have happened. 'It's almost simplistic,' he said. 'The imagination of the criminal exceeds that of the innocent.' His meaning was clear: a criminal can imagine committing his own crimes, while his victims cannot imagine *anyone* committing such crimes."⁶⁷

Because of something called the personal-positivity bias, people tend to make favorable rather than unfavorable judgments about others.⁶⁸ Even though people think that most politicians are crooks, they are much less likely to think that their own senators or governor are crooks.⁶⁹ Thus, investors will tend to expect, perhaps naively, that they

64. See Warren Mueller, *Residential Tenants and Their Leases: An Empirical Study*, 69 MICH. L. REV. 247, 274-75 (1970) (finding that tenants did not understand their lease contracts, but perceived that the terms were more favorable to them than they actually were).

65. See generally Colin Camerer, *Individual Decision Making*, in THE HANDBOOK OF EXPERIMENTAL ECONOMICS 587, 612-13 (John H. Kagel & Alvin E. Roth eds., 1995) (explaining the phenomenon).

66. See Donald C. Langevoort, *Where Were the Lawyers? A Behavioral Inquiry into Lawyers' Responsibility for Clients' Fraud*, 46 VAND. L. REV. 75, 107 (1993) (explaining how this effect induces lawyers with crooked clients to overlook obvious frauds); Prentice, *Irrational Auditor*, *supra* note 3, at 163 (noting how the false consensus effect can cause auditors with crooked clients to overlook obvious frauds).

67. DIANA B. HENRIQUES, *THE WIZARD OF LIES: BERNIE MADOFF AND THE DEATH OF TRUST* 213 (2011).

68. PLOUS, *supra* note 26, at 186 (noting studies finding that in some circumstances people will tend to attribute positive behaviors primarily to dispositional factors and negative behaviors primarily to situational factors); Allen Blair, *A Matter of Trust: Should No-Reliance Clauses Bar Claims for Fraudulent Inducement of Contract?*, 92 MARQ. L. REV. 423, 466 (2009) (noting this phenomenon).

69. Kathleen M. McGraw, *Manipulating Public Opinion with Moral Justification*, 560 ANNALS AM. ACAD. POL. & SOC. SCI. 129, 136-37 (1998); David O. Sears, *The Person-Positivity Bias*, 44 J. PERSONALITY & SOC. PSYCHOL. 233, 245

will be treated fairly. They know about Bernie Madoff, Alan Stanford, and other crooked securities industry professionals, but they are sure that the nice broker they have chosen will look out for their best interests.

D. Insensitivity to Source of Information

Another difficulty people have in avoiding being defrauded stems from their insensitivity to the source of information.⁷⁰ People often do not react sensibly even when their stock brokers give explanations for facts that seem to contradict their recommendations.⁷¹ Studies show that people have difficulty disregarding information, even when they learn that it is from an unreliable source.⁷² And they don't even realize it. In a recent study people who recognized that the sources of rumors were not credible and claimed that they were not influenced by those rumors actually traded stocks as if the rumors were real news.⁷³

An additional wrinkle here is that people generally believe that they are good at detecting when they are being lied to,⁷⁴ although the evidence is that most people are not.⁷⁵ Professor Donald Langevoort has argued that "once a broker successfully cultivates trust, willing

(1983) (finding evidence that people liked individual politicians and individual professors more than politicians as a whole or professors as a whole).

70. See generally Amos Tversky & Daniel Kahneman, *Judgment under Uncertainty: Heuristics and Biases*, in *JUDGMENT UNDER UNCERTAINTY: HEURISTICS AND BIASES* 3, 7-11 (Daniel Kahneman et al. eds., 1982).

71. Martha Brannigan, *Victims of Investment Scams Seem Condemned to Repeat Past Errors*, *WALL ST. J.*, Mar. 24, 1988 (noting the case of a fraud victim who "continued to invest even though the saleswoman offered only excuses for why his prior ventures hadn't started paying off").

72. See generally Tversky & Kahneman, *supra* note 70, at 7-11 (noting people's insensitivity to predictability; for example, if given a description of a company and asked to predict its profitability, people's predictions tend to remain the same regardless of whether the information is reliable or unreliable).

73. Nicholas DiFonzo & Prashant Bordia, *Rumor and Prediction: Making Sense (but Losing Dollars) in the Stock Market*, 71 *ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES* 329, 346 (1997) ("[R]umors do not have to be believed or trusted in order to powerfully affect trading, they simply have to make sense.").

74. See ARTHUR ALLEN LEFF, *SWINDLING AND SELLING* 84-87 (1976) (describing the difficulty people have in believing that they could be fooled).

75. EVELIN SULLIVAN, *THE CONCISE BOOK OF LYING* 206 (2001) ("In scientifically conducted experiments, the success rate of people being asked to sort out lies from truths, say by watching people on videotape either lying or telling the truth, has been shown to be poor."); Peter Vallentyne, *The Rationality of Keeping Agreements*, in *CONTRACTARIANISM AND RATIONAL CHOICE: ESSAYS ON DAVID GAUTHIER'S MORALS BY AGREEMENT* 177 (Peter Vallentyne ed., 1991) ("[I]n the real world, people's dispositions are opaque enough that it is often possible to deceive others into thinking that one is trustworthy.").

reliance by the sophisticated investor—imprudent though it may seem in hindsight—is quite likely and, for that reason alone, worthy of some protection.”⁷⁶ Relatedly, Professor Richard Shell has noted that “human perception overall is not a reliable defense to opportunistic behavior.”⁷⁷ For all these and other reasons, many and perhaps most customers who trust their brokers will tend to discount the relevance of an NFD disclosure and therefore remain vulnerable to mistreatment even after the disclosure.

E. Probabilities and Future Events

People generally are not especially skilled at calculating probabilities, substituting simple rule-of-thumb heuristics for statistical accuracy.⁷⁸ In particular, investors and others tend to underestimate low-probability risks,⁷⁹ even when those risks carry, as securities fraud does, the potential for great loss.⁸⁰ Studies show that consumers tend to excessively discount the long-term risk of product failure and often focus instead on the immediate gratification of a product carrying a lower price.⁸¹ Just as product failures occur relatively infrequently,⁸² investor frauds and instances of subordination of investor interests are arguably low-probability events that investors will tend to ignore even

76. Donald C. Langevoort, *Selling Hope, Selling Risk: Some Lessons for Law from Behavioral Economics about Stockbrokers and Sophisticated Customers*, 84 CALIF. L. REV. 627, 631 (1996).

77. G. Richard Shell, *Opportunism and Trust in the Negotiation of Commercial Contracts: Toward a New Cause of Action*, 44 VAND. L. REV. 221, 267 (1991).

78. See Paul Slovic & Sarah Lichtenstein, *Comparison of Bayesian and Regression Approaches to the Study of Information Processing in Judgment*, 6 ORGANIZATIONAL BEHAV. & HUM. PERFORMANCE 649, 724 (1971) (noting that people have a very difficult time weighting and combining information to make probabilistic decisions and therefore “resort to simplified decision strategies, many of which lead them to ignore or misuse relevant information”).

79. See Melvin Aron Eisenberg, *The Structure of Corporation Law*, 89 COLUM. L. REV. 1461, 1464–65 (1989) (noting that contracting parties tend to underestimate the likelihood of low-probability events occurring).

80. See Paul Slovic et al., *Regulation of Risk: A Psychological Perspective*, in REGULATORY POLICY AND THE SOCIAL SCIENCES 241, 260 (Roger G. Noll ed., 1985) (discussing decision making in the face of low probability events).

81. See William C. Whitford, *Comment on a Theory of Consumer Product Warranty*, 91 YALE L.J. 1371, 1383 (1982).

82. Michael I. Meyerson, *The Efficient Consumer Form Contract: Law and Economics Meets the Real World*, 24 GA. L. REV. 583, 599 (1990) (“[B]ecause most contracts do not result in any loss to consumers, and because consumers lack knowledge about the likelihood of any particular loss, consumers tend to treat the risk as too insubstantial to protect against; therefore, they do not read the contract, let alone attempt to negotiate terms.”).

when warned that their stock brokers are entitled under the law to subordinate their interests. In other contexts, courts and sellers realize that consumers often, without full appreciation of what they are doing, “assent to terms that effectively discard their legal rights.”⁸³ Investors are the same.⁸⁴ If they trust their broker, and therefore underestimate the likelihood of having their interests subordinated to that of the broker, investors will tend to insufficiently appreciate the danger of signing a contract containing an NFD disclaimer.

F. The Impact of Oral Communications

A written warning that a broker does not owe an investor a fiduciary duty is unlikely to have any significant impact on the investor’s decision making if elaborate (and even inconsistent) oral representations have already been made. Once people decide that they like a person enough to choose him or her as their broker (or dentist or doctor or lawyer), all their subsequent opinions of that broker’s words and actions will be substantially colored.⁸⁵ People whose success depends on the efforts of others tend naturally to form positive impressions of those on whom they depend.⁸⁶ This is sort of a mild version of the Stockholm syndrome.⁸⁷ Once this trusting relationship is established, people will tend to “overdraw” on the available information, concluding that they possess real knowledge about their broker’s future conduct.⁸⁸ They believe, probably mistakenly, that the snapshot of information that they have received when interviewing the

83. Michael I. Meyerson, *The Reunification of Contract Law: The Objective Theory of Consumer Form Contracts*, 47 U. MIAMI L. REV. 1263, 1301 (1993).

84. See G. Richard Shell, *Fair Play, Consent and Securities Arbitration: A Comment on Speidel*, 62 BROOK. L. REV. 1365, 1369 (1996). In the context of a form contract containing a mandatory arbitration clause, Shell notes that “[e]ven if customers do read the boilerplate [in a stock broker’s form contract], they are unlikely to focus on the seemingly remote contingency that they will someday want to sue their broker.” *Id.*

85. See Hanson & Kysar, *supra* note 51, at 1524 (“[C]onsumers can be significantly influenced by their perceptions of a seller’s conduct and intentions.”).

86. See Langevoort, *supra* note 66, at 99 (citing Steven L. Neuberg & Susan T. Fiske, *Motivational Influences on Impression Formation: Outcome Dependency, Accuracy-Driven Attention, and Individuating Processes*, 53 J. PERSONALITY & SOC. PSYCHOL. 431 (1987)).

87. The “Stockholm syndrome,” of course, is the notion that hostages will, as a survival mechanism, bond with their captors. Laura Fitzpatrick, *Stockholm Syndrome*, TIME, Sept. 14, 2009, at 19. Some variant of this concept can affect investors. Stuart Washington, *ASIC Hits the Jackpot*, AGE (Melbourne), Feb. 2, 2011, at 4 (noting that a strange Stockholm syndrome can develop around charismatic fundraisers).

88. Klas Borell, *Trust and Fraud: Occupation and Resistance in Norway, 1940-1945*, 27 J. POL. & MIL. SOC. 39, 40-41 (1999).

broker represents a reliable indicator of the broker's future long-term actions.

Once the investors decide to trust the broker, they will tend not to read seriously any written communications that might contain a fiduciary duty disclaimer. In daily life, oral communications tend to trump written communications.⁸⁹ Securities professionals such as stock brokers are aware of, and take advantage of, the fact that oral communications aimed at convincing potential investors of what a great guy or gal the broker is will often dominate inconsistent written documents.⁹⁰

All this trust is exacerbated by a likeability principle. Many stock brokers succeed in gaining customers by being likeable.⁹¹ People tend

89. Robert Shiller notes:

The channels of human communication that we know today seem to favor the interpersonal face-to-face and word-of-mouth communication that developed over millions of years of evolution, during times when such communication was virtually the only form of interpersonal communication. The patterns of communication hard-wired into our brains rely on there being another person's voice, another person's facial expressions, another person's emotions, and an associated environment of trust, loyalty, and cooperation. Because these elements are missing from the written or electronic word, people find it somewhat more difficult to react to these sources of information. They cannot give these other sources the same emotional weight, nor can they remember or use information from these other sources as well. This is an important reason why we still have teachers—why we cannot tell our children simply to sit down and read books or rely on computer-aided instruction.

ROBERT J. SHILLER, *IRRATIONAL EXUBERANCE* 155 (2000).

90. In counseling stock brokers on how to effectively market securities, Steven Drozdeck and Karl Gretz note:

According to Albert Mehrabian, 55 percent of our communication is through posture, gestures, and facial expressions; 38 percent is through our tone of voice; and only 7 percent through words themselves. These percentages have proven themselves valid through scores of other studies that also indicate that 85 to 93 percent of communication occurs below the conscious level of awareness. Therefore, in a face-to-face presentation you are able to deliver 100 percent of the message, while a telephone presentation allows you to deliver only 45 percent; and a letter only 7 percent.

STEVEN R. DROZDECK & KARL F. GRETZ, *THE BROKER'S EDGE: HOW TO SELL SECURITIES IN ANY MARKET* 222 (1995) (footnote omitted); see also Donald C. Langevoort, *Toward More Effective Risk Disclosure for Technology-Enhanced Investing*, 75 WASH. U. L.Q. 753, 761-62 (1997) ("[T]he permissibility of oral selling efforts during the waiting period invites promoters to persuade buyers of the virtue of the investment before there is much of an opportunity for review of the required disclosures . . . [so that] salesmanship can readily trump the late-arriving prospectus (if the investor ever had any inclination to read it at all)." (footnotes omitted)).

91. "How-to" books on being a successful stock broker often emphasize how to be likeable. See DROZDECK & GRETZ, *supra* note 90, at 37-38.

to like others who are like them, who compliment them, and who are physically attractive. Tupperware parties succeed on the same principle, by setting up a situation where the seller and potential buyers are friends.⁹² The enduring success of “affinity scams” stems from this principle.⁹³ Even sophisticated investors often fall victim to such schemes.⁹⁴ Once stock brokers have convinced their customers that they are friends, the impact of any NFD warning is likely to be largely nullified.

G. Social Proof

In most settings, people tend to take their cues as to proper behavior from those around them. This is often called “social proof.”⁹⁵ Marketers sensibly take advantage of this phenomenon by representing products as the “best-selling” or “fastest-growing.” Regarding standard arbitration clauses that take away the right to sue, Professor Shell has noted:

To the extent customers think about the arbitration clause [in a broker’s form contract] at all, my guess is that they say to themselves: “There are a lot of people like me investing in the stock market and they all signed this contract, too. Some of them must have given this some thought even if I am too busy to do so. It must be OK.”⁹⁶

This same line of reasoning is quite likely to lead people to accept without undue disquiet an NFD clause. They are likely to accept it automatically without considering its ramifications, especially since no obvious alternative options are likely to present themselves.

92. See Rex Taylor, *Marilyn’s Friends and Rita’s Customers: A Study of Party-Selling as Play and as Work*, 26 SOC. REV. 573, 590-92 (1978) (studying the dynamics of such selling methods).

93. Affinity schemes arise when fraudsters use common religious, ethnic, or political affiliations with investors to gain their trust and then manipulate it for selfish purposes. Mark A. Sargent, *A Future for Blue Sky Law*, 62 U. CIN. L. REV. 471, 509 (1993) (discussing affinity schemes).

94. Of course, Bernie Madoff’s ripping off of the Jewish community was clearly an affinity scheme within a larger fraud. See Christine Hurt, *Evil Has a New Name (and a New Narrative): Bernard Madoff*, 2009 MICH. ST. L. REV. 947, 957-59; Felicia Smith, *Madoff Ponzi Scheme Exposes “The Myth of the Sophisticated Investor”*, 40 U. BALT. L. REV. 215, 228-30 (2010).

95. ROBERT B. CIALDINI, *INFLUENCE: SCIENCE AND PRACTICE* 95 (3d ed. 1993).

96. Shell, *supra* note 84, at 1371.

H. Difficulty of Judgment Correction

Once people have developed a positive enough impression of people to want to hire them as their stock brokers, it will be difficult for them to change their minds on the basis of new information, particularly information that is no more useful than an "I, like every other stock broker, have a conflict of interest" disclosure.

Daylian Cain and colleagues point out that people's ability to undo a biasing influence on their own judgment is severely limited.⁹⁷ To do so requires three preconditions: (1) motivation "to make correct judgments," (2) awareness of "potentially distorting influence[s]," and (3) awareness of "the direction *and* magnitude of [the distorting] influences."⁹⁸

1. MOTIVATION TO MAKE CORRECT JUDGMENTS

One might think that people would want to make correct judgments, but at some level this is not necessarily true. People like to act consistently,⁹⁹ for this is generally viewed as logical, rational, and honest conduct.¹⁰⁰ Political campaigns illustrate that politicians who act inconsistently with previous statements or positions are often heavily punished by voters.¹⁰¹ Politicians often stick with positions that are obviously wrong for the sole purpose of maintaining an appearance of consistency, hobgoblins of small minds notwithstanding.¹⁰²

Stock brokers have played upon this desire for consistency in order to induce investors to make purchases they did not necessarily want or need.¹⁰³ Once investors, even sophisticated investors, choose a stock

97. Daylian M. Cain et al., *Coming Clean but Playing Dirtier: The Shortcomings of Disclosure as a Solution to Conflicts of Interest*, in *CONFLICTS OF INTEREST: CHALLENGES AND SOLUTIONS IN BUSINESS, LAW, MEDICINE, AND PUBLIC POLICY* 104, 109 (Don A. Moore et al. eds., 2005).

98. *Id.*

99. CIALDINI, *supra* note 95, at 51 ("Once we make a choice or take a stand, we will encounter personal and interpersonal pressures to behave consistently with that commitment." (emphasis omitted)).

100. *Id.* at 52.

101. Few political candidates have been punished more for flip-flopping than presidential candidate John Kerry, who was for the Iraq war before he was against it. See Holning Lau, *Identity Scripts & Democratic Deliberation*, 94 MINN. L. REV. 897, 915 (2010).

102. See Rebecca M. Kysar, *Listening to Congress: Earmark Rules and Statutory Interpretation*, 94 CORNELL L. REV. 519, 531 (2009) (noting that increased transparency makes it harder for politicians to alter their positions due to danger of being labeled a flip-flopper).

103. A classic example (in fact, several) occurred during the Bernie Madoff Ponzi scheme. French fund manager Thierry de la Villehuchet was warned repeatedly

broker, then for consistency reasons they will often wish strongly to trust that broker regardless of formal warnings that the broker does not owe them a duty to put their interests ahead of his or her own.¹⁰⁴

The cognitive dissonance phenomenon further reinforces these tendencies. It tends to cause investors who have placed confidence in a seller of securities to be extremely reluctant to reach the conclusion that they made a mistake in reposing that confidence even when contrary evidence begins to come to light.¹⁰⁵ Instead, they will tend to see what they want to see.¹⁰⁶ Cognitive dissonance tends to cause physicians who desire to continue to receive wide-ranging benefits from drug companies to believe that their judgments are not affected by the lavish gifts.¹⁰⁷

and in very explicit terms that Madoff's investment scheme had to be a Ponzi scheme, but because he was "totally committed" to Madoff, with other people's money, he felt he just had to trust Madoff. He committed suicide after the scandal came to light. HARRY MARKOPOLOS, *NO ONE WOULD LISTEN* 91, 218 (2010).

104. Professor Langevoort notes:

Most investors lack the time or information necessary to digest all information relating to potential investments, yet they feel substantial pressure to act. They anticipate the sort of regret that comes when they have no one to blame for making unwise choices but themselves, and thus they are highly motivated to seek out apparent experts, such as stockbrokers and investment advisers, shifting away the locus of responsibility. Having committed to a particular expert, of course, investors feel a strong tendency to bolster their choice. *Under these circumstances, reliance is apt to occur regardless of formal warnings or disclosures. To heed them against the recommendation of one's chosen broker or adviser is to admit the possibility that one has chosen foolishly.*

Donald C. Langevoort, *Ego, Human Behavior, and Law*, 81 VA. L. REV. 853, 879 (1995) (emphasis added) (footnotes omitted).

105. Cognitive dissonance is the tendency of people to avoid or minimize psychological inconsistencies. *See, e.g.*, LEON FESTINGER, *A THEORY OF COGNITIVE DISSONANCE* 2-3 (1957); MARGARET HEFFERNAN, *WILLFUL BLINDNESS* 51-57 (2011); Lisa L. Shu et al., *When to Sign on the Dotted Line?* 8 (Harvard Bus. Sch., Working Paper No. 11-117, 2011), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1837611. Thus, someone who has made a particular decision will thereafter tend to look for evidence that supports the decision and tend to ignore evidence that undermines it. Indeed, people whose beliefs are challenged will tend to be even more vociferous in their espousal of the beliefs than people whose beliefs are not challenged. *See* David Gal & Derek D. Rucker, *When in Doubt, Shout!: Paradoxical Influences of Doubt on Proselytizing*, *PSYCHOL. SCI.* (forthcoming Nov. 2010) (proof at 6), available at <http://pss.sagepub.com/content/early/2010/10/12/0956797610385953>.

106. *See* Ripken, *supra* note 17, at 176; Vandenberg et al., *supra* note 26, at 759 (noting that people often reduce dissonance through rationalization).

107. Susan Chimonas et al., *Physicians and Drug Representatives: Exploring the Dynamics of the Relationship*, 22 J. GEN. INTERNAL MED. 184, 189 (2007) (concluding that physicians' techniques, including massive rationalization, for handling cognitive dissonance render it unlikely that physicians would police themselves).

The simple fact is that once people believe X, it is difficult for them to believe Y and they may not wish to do so. People's brains treat differently information that supports their existing beliefs and conclusions than it does information that contradicts them.¹⁰⁸ Drew Westen and his colleagues performed fMRI scans on the brains of people as they read quotations attributed either to the presidential candidate that they supported or the one that they opposed. The confirming information activated reward circuits in the brain whereas the contradictory information created distress and the brain quickly shut that distress down via faulty reasoning.¹⁰⁹ People's brains reward the intake of confirming information and struggle with processing contradictory information,¹¹⁰ often making it easier to remain consistent than to change to a more accurate belief.

2. AWARENESS OF POTENTIAL DISTORTING INFLUENCES

Most people understand the potentially deforming effects of a conflict of interest well enough that they are aware of its potentially distorting influence. They will comprehend that the advice of a conflicted agent is probably not as trustworthy as that of an unconflicted agent. Therefore, this second requirement needed to undo a biasing influence would typically be met in the case of an NFD disclosure, unlike the first or the third.

3. AWARENESS OF THE DIRECTION AND MAGNITUDE OF THE DISTORTING INFLUENCES

Any revision of judgment taking into account new evidence of a conflict of interest is likely to be insufficient for several reasons. The first is the confirmation bias.¹¹¹ Because of this bias, people tend to look for information that supports previous conclusions and to overlook inconsistent information. Therefore, customers who have hired brokers

108. HEFFERNAN, *supra* note 105, at 44.

109. DREW WESTEN, *THE POLITICAL BRAIN: THE ROLE OF EMOTION IN DECIDING THE FATE OF THE NATION* xiii-xiv (2007).

110. *Id.*

111. People tend to look for information to confirm decisions they have already made, overweigh evidence that confirms previous conclusions, and to underweigh evidence that undermines it. This is the confirmation bias. *See, e.g.*, PAUL BREST & LINDA HAMILTON KRIEGER, *PROBLEM SOLVING, DECISION MAKING, AND PROFESSIONAL JUDGMENT: A GUIDE FOR LAWYERS AND POLICY MAKERS* 609, 615, 618-19 (2010); Ryan Rodenberg, *Perception ≠ Reality: Analyzing Specific Allegations of NBA Referee Bias*, 7 J. QUANTITATIVE ANALYSIS SPORTS, issue 2, art. 8, 2011 at 1, 7-8 (noting that confirmation bias may help explain NBA coaches' unfounded beliefs that certain referees are biased against their teams).

based upon favorable impressions of them will tend to look for information that supports the judgment they have made (“this broker will look out for my best interests”) and ignore evidence that contradicts that judgment.¹¹²

A second reason that the adjustment in judgment will likely be insufficient arises from the anchoring heuristic.¹¹³ People tend to anchor on the first information they receive, and then revise their judgments in the face of new information, but to an insufficient degree.¹¹⁴ Thus, it has been suggested (by the author) that if a customer is given plentiful oral representations by a stock broker regarding how great a stock is and how well it will perform, he or she may anchor on those promises and not adjust adequately when given information in a written contract disclaiming any representations or warranties.¹¹⁵ Similarly, if customers form a favorable impression of a broker (“this broker is friendly and nice”) and hire him or her, then even if they recognize that subsequent information is inconsistent with the earlier information, they have a strong tendency to take the new information into account, if at all, to an insufficient degree. As Robertson observes, “once a layperson-advisor relationship is created, it is likely to be sticky. . . . Once the layperson has a relationship with a conflicted advisor, a mere disclosure may not be enough to break that connection.”¹¹⁶

As an illustration, Cain and colleagues performed an experiment wherein participants were asked a question and received advice regarding the right answer from another person. Some of the advisors

112. Anders Kaye, *Does Situationist Psychology Have Radical Implications for Criminal Responsibility?*, 59 ALA. L. REV. 611, 637–38 (2008) (noting that the confirmation bias tends to cause people to reject information that conflicts with their working models of people and “that once we have decided [that someone] is a certain ‘kind’ of person,” it is difficult to change that opinion).

113. The essential problem with anchoring and adjustment is that people often anchor on original information and then adjust insufficiently when new, inconsistent information comes to light. *See generally*, BREST & KRIEGER, *supra* note 111, at 267; Amos Tversky & Daniel Kahneman, *Judgment under Uncertainty: Heuristics and Biases*, 185 SCIENCE 1124, 1128 (1974); Timothy E. Lynch, *Gambling by Another Name? The Challenge of Purely Speculative Derivatives* 21 (Ind. Univ. Maurer Sch. of Law-Bloomington Legal Studies Research Paper Series, Research Paper No. 188, 2011), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1788219.

114. Ward Edwards, *Conservatism in Human Information Processing*, in JUDGMENT UNDER UNCERTAINTY *supra* note 70, at 359, 359 (“[O]pinion change is very orderly, and usually proportional to numbers calculated from Bayes’s theorem—but it is insufficient in amount.”); Joseph F. Funaro, *An Empirical Analysis of Five Descriptive Models for Cascaded Inference*, 14 ORGANIZATIONAL BEHAV. & HUM. PERFORMANCE 186, 186 (1975). Most empirical evidence indicates that “intuitive opinion revisions are conservative in comparison to the optimal revisions specified by Bayes’ theorem.” *Id.*

115. Prentice, *Contract-Based Defenses*, *supra* note 3, at 374.

116. Robertson, *supra* note 28, at 695.

were compensated by how accurate the participants' guesses were; others were compensated by how *high* the participants' guesses were. Participants in the "high" condition were specifically warned: "When you answer, remember that the suggested answers were provided by someone who was trying to get you to give an answer that was artificially high."¹¹⁷ Despite that warning, the subjects anchored heavily upon the suggested answer and did not move far away from it. "[O]verall test scores were driven largely by the anchor suggested to them [And t]he disclosures [regarding conflict of interest] had no significant main or interaction effect on participants' responses."¹¹⁸

To learn of a stock broker's conflict of interest certainly signals a potential need to revise a previous conclusion regarding the broker's objectivity, but it remains nearly impossible for an investor to know how much to revise the previous conclusion, especially in light of the fact that other brokers similarly owe no fiduciary duty.¹¹⁹

The factors mentioned in the preceding discussion are but the tip of the iceberg regarding the vast array of data indicating that the manner in which people gather and process information, even when it is put right in front of them in readily digestible form, leaves them quite vulnerable to fraud and other forms of mistreatment. One who studies the various heuristics and biases to which humans are subject can quickly become overwhelmed by the vast number of them. Even a partially comprehensive article on the topic quickly produces a lengthy laundry list of limitations upon rationality. It has been suggested that such a list may go too far,¹²⁰ but each year more evidence undermining the rational man model of decision making piles up, and the numerous studies recently surveyed by Professors Omri Ben-Shahar and Carl Schneider regarding the failures of mandated disclosure¹²¹ provide persuasive support for the positions taken in this Article. As they conclude, "mandated disclosure is fundamentally misconceived because its solution to the problem of choice is information alone. But people's problems choosing go well beyond ignorance."¹²²

117. Cain et al., *supra* note 97, at 111.

118. *Id.* at 111-12.

119. *See id.* at 112.

120. *See, e.g.*, Gregory Mitchell, *Taking Behavioralism Too Seriously? The Unwarranted Pessimism of the New Behavioral Analysis of Law*, 43 WM. & MARY L. REV. 1907, 1911-13 (2002) (suggesting that proponents of behavioral law and economics overpromise); Gregory Mitchell, *Why Law and Economics' Perfect Rationality Should Not Be Traded for Behavioral Law and Economics' Equal Incompetence*, 91 GEO. L.J. 67 (2002) (similar).

121. Ben-Shahar & Schneider, *supra* note 47, at 653-65.

122. *Id.* at 720.

III. BOUNDED ETHICALITY, MORAL LICENSE, AND THE LIMITATIONS OF DISCLOSURE

As the previous Section indicated, cognitive limitations and behavioral heuristics often prevent investors from adequately protecting themselves even in the presence of full and candid disclosure. Mandated disclosure also often fails because of problems on the discloser's side of the equation.¹²³ In some settings, required disclosure embarrasses a firm or individual so much that they feel constrained to avoid unsavory activity.¹²⁴ However, that is often not the case. For purposes of this Article, the more important concepts are emerging from the new field of behavioral ethics—bounded ethicality, moral equilibrium, and their consequences.¹²⁵

A. Bounded Ethicality

As noted above, various cognitive and behavioral forces render people boundedly rational. They may also cause people to be boundedly ethical, to make moral mistakes even though they are well intentioned.¹²⁶ In one of the most recent and fullest accounts of bounded ethicality, Dolly Chugh, Max Bazerman, and Mahzarin Banaji¹²⁷ integrated (a) Herbert Simon's notion of bounded rationality,¹²⁸ (b) the Kahneman and Tversky-inspired heuristics and biases literature,¹²⁹ and

123. Individuals or firms required to disclose financial or other data often have difficulty interpreting the legislative or regulatory mandate, assembling the data, and effectively implementing the mandate (even if they desire to do so). *Id.* at 692–93. Commonly, they will resist good faith compliance with the disclosure mandate and thereby defeat its best intentions. *Id.* at 698–702.

124. FUNG ET AL., *supra* note 21, at 29 (giving an example of 1986 requirements for disclosure of toxic pollution released into the environment that caused firms to clean up their acts).

125. See Ming Li & Kristóf Madarász, *When Mandatory Disclosure Hurts: Expert Advice and Conflicting Interests*, 139 J. ECON. THEORY 47 (2008) (using economic models, rather than the behavioral principles emphasized in this Article, to identify contexts where mandatory disclosure arguably prevents meaningful communication).

126. “[B]ounded ethicality is a strand [of the broader concept of bounded rationality] that is used to refer to the limits on the quality of decision making with ethical import.” Dolly Chugh et al., *Bounded Ethicality as a Psychological Barrier to Recognizing Conflicts of Interest*, in CONFLICTS OF INTEREST, *supra* note 97, at 74, 75.

127. *Id.* at 74.

128. See *supra* notes 44–46 and accompanying text.

129. See generally CHOICES, VALUES, AND FRAMES (Daniel Kahneman & Amos Tversky eds., 2000); HEURISTICS AND BIASES: THE PSYCHOLOGY OF INTUITIVE JUDGMENT (Thomas Gilovich et al. eds., 2002).

(c) recent insights regarding the limitations of the conscious mind,¹³⁰ to conclude that “these three literatures together provide robust support for the view that conflict of interest is not limited to explicit dishonesty. Rather, unconscious acts of ethically questionable behavior are more prevalent, more insidious, and as such, more in need of attention.”¹³¹

The most discomfoting thing about bounded ethicality is that well-intentioned people can make serious ethical errors without ever consciously deciding to stray from the straight and narrow.¹³² For example, advisors like stock brokers can and often do give biased advice without consciously realizing that it is biased.¹³³ How does this happen?

It is now taken as a given that human decision making occurs along two tracks—both unconscious and conscious.

[T]he human mind has two processing systems that can operate simultaneously. What is often termed “System 1” is an intuitive system that works automatically, rapidly, and effortlessly. We often use this system without even realizing it. System 2 is a reflective system that is controlled, effortful, deductive, and self-aware. In the anchoring and adjustment phenomenon, it is System 1 that automatically anchors on a number given, and System 2 that consciously adjusts away from that anchor (but usually not far enough).¹³⁴

130. See, e.g., JONATHAN HAIDT, *THE HAPPINESS HYPOTHESIS* 165 (2006). Haidt argues that emotional responses underlie many moral judgments and that “[t]rying to make children behave ethically by teaching them to reason well is like trying to make a dog happy by wagging its tail. It gets causality backwards.” *Id.*

131. Chugh et al., *supra* note 126, at 75.

132. See Max H. Bazerman & Mahzarin R. Banaji, *The Social Psychology of Ordinary Ethical Failures*, 17 *SOC. JUST. RES.* 111, 111 (2004) (“[W]e believe that . . . the vast majority of unethical behaviors . . . occur without the conscious awareness of the actors who engage in them.”).

133. See Don A. Moore et al., *Conflicts of Interest and the Case of Auditor Independence: Moral Seduction and Strategic Issue Cycling*, 31 *ACAD. MGMT. REV.* 10, 11 (2006) (“We argue . . . that doctors’ advice is biased . . . and that they typically believe their biased advice is unbiased.”); Robertson, *supra* note 28, at 680 (“Advisors may not even know that they are biased by their conflicting interests.”).

134. Prentice, *Chicago Man*, *supra* note 3, at 1717. See generally BROOKS, *supra* note 55, at x-xi (discussing dual-processing); KEITH E. STANOVICH, *WHO IS RATIONAL? STUDIES OF INDIVIDUAL DIFFERENCES IN REASONING* 144-47 (1999) (introducing a dual-processing model of decision making); Anna Ronkainen, *Dual-Process Cognition and Legal Reasoning* 1-2 (June 30, 2011) (unpublished working paper), available at <http://ssrn.com/abstract=1879429> (analyzing dual-process cognition in context of legal reasoning).

Recent studies have found that System 1's unconscious impact on decision making is extensive and significant.¹³⁵ Although it is difficult for many people to accept, there is substantial evidence "that most of a person's everyday life is determined not by their conscious intentions and deliberate choices but by mental processes that are put into motion by features of the environment and that operate outside of conscious awareness and guidance."¹³⁶

One of the most powerful unconscious forces affecting people's judgment is the self-serving bias. The self-serving bias is a pervasive and subconscious influencer of human judgments and actions.¹³⁷ For both cognitive and motivational reasons, all people are self-serving in that their judgments and actions will tend to favor whatever is in their own best interests or whatever is consistent with previous positions they have taken.¹³⁸

Although the self-serving bias obviously will not dominate for all people in all settings, it tends to affect how people seek out

135. See generally Jonathan St. B. T. Evans, *In Two Minds: Dual-Process Accounts of Reasoning*, 7 TRENDS COGNITIVE SCI. 454 (2003); Steven A. Sloman, *Two Systems of Reasoning*, in HEURISTICS AND BIASES, *supra* note 129, at 379, 379-96. This phenomenon has begun to get serious notice in the legal literature. See, e.g., Kevin Emerson Collins, *Constructive Nonvolition in Patent Law and the Problem of Insufficient Thought Control*, 2007 WIS. L. REV. 759, 795-96 (discussing the acts of thinking that can fall within a patent claim); John F. Irwin & Daniel L. Real, *Unconscious Influences on Judicial Decision-Making: The Illusion of Objectivity*, 42 MCGEORGE L. REV. 1, 1-2 (2010) (discussing the impact of unconscious bias upon judicial decision making); Vandenberg et al., *supra* note 26, at 758 ("Psychologists have hypothesized that many 'decisions' actually occur outside the range of cognition."); W. Bradley Wendel, *The Behavioral Psychology of Judicial Corruption: A Response to Judge Irwin and Daniel Real*, 42 MCGEORGE L. REV. 1 (2010) (discussing the impact of unconscious bias upon judicial decision making).

136. John A. Bargh & Tanya L. Chartrand, *The Unbearable Automaticity of Being*, 54 AM. PSYCHOLOGIST 462, 462 (1999). See generally TIMOTHY D. WILSON, STRANGERS TO OURSELVES: DISCOVERING THE ADAPTIVE UNCONSCIOUS 14 (2002) (noting the overwhelmingly significant role the unconscious plays in human behavior and decision making).

137. See generally BREST & KRIEGER, *supra* note 111, at 244-49 (discussing a series of impacts of the self-serving bias); Prentice, *SEC and MDP*, *supra* note 3, at 1604-53 (explaining the self-serving bias generally, but focusing on its implications for auditors).

138. The self-serving bias's existence and impact are settled, but its causes are the subject of debate. It seems clear that both cognitive and motivational factors play a role. See generally Tom Pyszczynski et al., *Maintaining Consistency between Self-Serving Beliefs and Available Data: A Bias in Information Evaluation*, 11 PERSONALITY & SOC. PSYCHOL. BULL. 179, 186-88 (1985) (concluding that both cognitive and motivational forces accounted for self-serving results in a study).

information,¹³⁹ perceive information,¹⁴⁰ process and encode information,¹⁴¹ and even how they recall information.¹⁴² Evidence demonstrates that the self-serving bias affects virtually every category of professional,¹⁴³ including lawyers,¹⁴⁴ physicians,¹⁴⁵ auditors,¹⁴⁶

139. Studies of human eye movements indicate that we studiously avoid even seeing things that may discomfort us. Lester Luborsky et al., *Looking, Recalling, and GSR as a Function of Defense*, 70 J. ABNORMAL PSYCH. 270, 276 (1965).

140. One of the most famous studies of the self-serving bias indicated that fans of two teams at a football game are each likely to come away from the game perceiving that the referees favored the other team. Albert H. Hastorf & Hadley Cantril, *They Saw a Game: A Case Study*, 49 J. ABNORMAL & SOC. PSYCHOL. 129, 131-32 (1954). Similarly, proponents of two political candidates are each likely to perceive that the candidate they favor prevailed in a political debate. David O. Sears & Richard E. Whitney, *Political Persuasion*, in HANDBOOK OF COMMUNICATION 253, 255 (Ithiel de Sola Pool et al. eds., 1973).

141. Charles G. Lord et al., *Biased Assimilation and Attitude Polarization: The Effects of Prior Theories on Subsequently Considered Evidence*, 37 J. PERSONALITY & SOC. PSYCHOL. 2098, 2105 (1979) (finding that when opponents and proponents of an issue are given two opposing studies, they tend to find the one that favors their position to be more persuasive); Leigh Thompson & George Loewenstein, *Egocentric Interpretations of Fairness and Interpersonal Conflict*, 51 ORG. BEHAV. & HUM. DECISION PROCESSES 176, 180-81 (1992) (finding that people tend to perceive information that favors their position as more important than information they run across that does not).

142. Baruch Fischhoff & Ruth Beyth, *"I Knew It Would Happen": Remembered Probabilities of Once-Future Things*, 13 J. ORG. BEHAV. & HUM. PERFORMANCE 1, 7 (1975) (finding that students remembered having made better predictions than they actually did); Thompson & Loewenstein, *supra* note 141, at 189-90 (finding that negotiators tend to remember facts that supported their position better than facts that undermined it).

143. For an informative and up-to-date collection of papers discussing the broad range of conflict of interest embodiments of the self-serving bias, see CONFLICTS OF INTEREST, *supra* note 97 (containing sections on business, medicine, law, and public policy).

144. Samuel Issacharoff, *Legal Responses to Conflicts of Interest*, in CONFLICTS OF INTEREST, *supra* note 97, at 189, 200 ("Conflicts of interest are endemic in the law."); Donald C. Langevoort & Robert K. Rasmussen, *Skewing the Results: The Role of Lawyers in Transmitting Legal Rules*, 5 S. CAL. INTERDISC. L.J. 375, 437-38 (1997) (suggesting that the self-serving bias is likely to lead attorneys to skew advice to clients in such a way as to increase the attorneys' seeming importance to the client).

145. See, e.g., HEFFERNAN, *supra* note 105, at 185 (quoting orthopedic surgeon David Ring as saying that "[d]octors who own stakes in testing labs order more tests"); Thomas L. Carson, *Conflicts of Interest*, 13 J. BUS. ETHICS 387, 394 (1994) ("[W]hen physicians are paid according to how much work they do for their patients, many physicians succumb to the temptation to provide their patients with unnecessary, even dangerous treatments."); Jerome P. Kassirer, *Physicians' Financial Ties with the Pharmaceutical Industry*, in CONFLICTS OF INTEREST, *supra* note 97, at 133, 134 (detailing tight web of conflicts of interests between pharmaceutical firms and physicians); Andrew Stark, *Why Are (Some) Conflicts of Interest in Medicine So Uniquely Vexing?*, in CONFLICTS OF INTEREST, *supra* note 97, at 152 (looking at conflicts of interest for both academic researchers and practicing physicians).

investment bankers,¹⁴⁷ securities analysts,¹⁴⁸ scientists,¹⁴⁹ policy researchers,¹⁵⁰ expert witnesses,¹⁵¹ judges,¹⁵² and, of course, stock brokers.¹⁵³ Importantly for present purposes, the self-serving bias

146. See David B. Citron & Richard J. Taffler, *The Audit Report under Going Concern Uncertainties: An Empirical Analysis*, 22 ACCT. & BUS. RES. 337, 344 (1992) (finding that auditors tended to issue going concern qualifications only to clients that were both in a very weak financial condition and in imminent danger of failing); Michael K. Shaub & Janice E. Lawrence, *Ethics, Experience and Professional Skepticism: A Situational Analysis*, 8 BEHAV. RES. ACCT. 124, 155 (Supp. 1996) (“[A]uditors’ professional skepticism was . . . counteracted when the client was important to the audit firm’s practice development.”).

147. Steven J. Cleveland, *An Economic and Behavioral Analysis of Investment Bankers When Delivering Fairness Opinions*, 58 ALA. L. REV. 299, 330 (2006) (noting the significant impact of the self-serving bias upon the actions and decisions of investment bankers); Robert A. Prentice & John H. Langmore, *Hostile Tender Offers and the “Nancy Reagan Defense”: May Target Boards “Just Say No”? Should They Be Allowed To?*, 15 DEL. J. CORP. L. 377, 468–69 (1990) (citing sources indicating that self-interest in the form of potential fees from clients sway investment bankers’ fairness opinions). Judges from time to time slam investment banks for their conflict-driven self-dealing. See, e.g., *In re Del Monte Foods Co. Shareholders Litig.*, No. 6027-VCL, 2011 Del. Ch. LEXIS 94 (Del. Ch. Feb. 14, 2011) (criticizing Barclay’s Bank for its actions in the Del Monte takeover case).

148. Mathew L.A. Hayward & Warren Boeker, *Power and Conflicts of Interest in Professional Firms: Evidence from Investment Banking*, 43 ADMIN. SCI. Q. 1, 14–15 (1998) (finding substantial evidence that an investment banking firm’s self-interest in the form of potential fees causes its stock analysts to rate more highly the securities of its clients than do other stock analysts).

149. See, e.g., Henry Thomas Stelfox et al., *Conflict of Interest in the Debate Over Calcium-channel Antagonists*, 338 NEW ENG. J. MED. 101, 103 (1998) (finding that ninety-six percent of a controversial drug’s defenders in the scientific community had financial ties to the drug’s maker).

150. See, e.g., Robert J. MacCoun, *Conflicts of Interest in Public Policy Research*, in CONFLICTS OF INTEREST, *supra* note 97, at 233, 234–35 (noting numerous examples of self-serving conflicts of interest involving policy researchers).

151. See, e.g., Paul S. Applebaum, *In the Wake of Ake: The Ethics of Expert Testimony in an Advocate’s World*, 15 BULL. AM. ACAD. PSYCHIATRY & L. 15, 21 (1987) (observing the “frequency with which highly respected [psychiatric] experts arrive at conclusions favorable to the side for which they are working or to which they have been assigned”).

152. See Frank B. Cross & Emerson H. Tiller, *Judicial Partisanship and Obedience to Legal Doctrine: Whistleblowing on the Federal Courts of Appeals*, 107 YALE L.J. 2155, 2175 (1998) (reporting an empirical study finding that D.C. Circuit Court three-judge panels controlled by Republican-appointed judges were more likely to defer to conservative administrative agency decisions than were panels controlled by Democrat-appointed judges); David B. Spence & Paula Murray, *The Law, Economics, and Politics of Federal Preemption Jurisprudence: A Quantitative Analysis*, 87 CALIF. L. REV. 1125, 1195 (1999) (finding in empirical study that decisions about federal preemption in environmental cases are the result of “actions of (partly) ideologically-motivated federal judges”).

153. Prentice, *Contract-Based Defenses*, *supra* note 3, at 399–400 (discussing implications of the self-serving bias for stock brokers and others).

affects how people judge the fairness of situations and of their own actions.¹⁵⁴

Given the abiding self-serving bias, as well as the overconfidence bias mentioned earlier in this Article,¹⁵⁵ it is not surprising that automatic decision making tends to unconsciously favor the self.

Ethical decisions are biased by a stubborn view of oneself as moral, competent, and deserving, and thus, not susceptible to conflicts of interest. To the self, a view of morality ensures that the decision maker resists temptations for unfair gain; a view of competence ensures that the decision maker qualifies for the role at hand; and, a view of deservingness ensures that one's advantages arise from one's merits. An ethical blind spot emerges as decision makers view themselves as moral, competent, and deserving, and thus assume that conflicts of interest are nonissues.¹⁵⁶

Thus, General James Marks (ret.), who while serving as an expert military analyst on CNN was discovered to have been simultaneously serving as president of a company seeking defense contracts in Iraq said, and no doubt believed: "I've got zero challenge separating myself from a business interest."¹⁵⁷ This overconfidence is made even more worrisome by Jonathan Haidt's theory (and evidence) that many times our moral decisions are made primarily (and quickly) by System 1, while System 2's involvement is primarily for the purpose of creating rationalizations for the decisions already made by System 1.¹⁵⁸

In the specific conflict of interest setting at stake in this article, Dolly Chugh and colleagues conclude:

[P]erceptual, cognitive, and social cognitive processes are bounded in similar, systematic ways that lead to gaps in observation and errors in decision making. Despite this robust

154. See Christine Jolls et al., *A Behavioral Approach to Law and Economics*, 50 STAN. L. REV. 1471, 1501 (1998) ("[P]arties may tend to see things in the light most favorable to them; while people care about fairness, their assessments of fairness are distorted by their own self-interest.").

155. See *supra* notes 55-59 and accompanying text.

156. Chugh et al., *supra* note 126, at 80.

157. David Barstow, *Behind TV Analysts, Pentagon's Hidden Hand*, N.Y. TIMES, Apr. 20, 2008, at A1.

158. See generally Jonathan Haidt, *The Emotional Dog and Its Rational Tail: A Social Intuitionist Approach to Moral Judgment*, 108 PSYCHOL. REV. 814, 830 (2001) (arguing for a "social intuitionist model" in favor of a rationalist model of ethical decision making and concluding, based on much new scientific evidence "that moral emotions and intuitions drive moral reasoning, just as surely as a dog wags its tail").

evidence about boundedness, humans tend to view their own ethicality as unbounded. In fact, decision makers are psychologically motivated to maintain a stable view of a self that is moral, competent, and deserving, and thus, immune from ethical challenges. Because individuals view their immunity as more powerful than the situation (moral, competent) and view any gains incurred as appropriate (competent, deserving), this view is a barrier to recognizing and addressing conflicts of interest. So, ironically, decision makers' persistent view of their own ethicality leads to subethical decisions.¹⁵⁹

Consider this as a plausible explanation for the continued existence of whole life insurance policies—in a setting where virtually everyone agrees that they are suboptimal for the consumer but they carry greater compensation for the insurance agents¹⁶⁰—and for much of the subprime mess. As prominent psychology professor John Darley noted:

Many decisions made intuitively are reasonable. . . . But the problem is that sometimes they are not. And the reasoning system monitoring—System Two—is intermittent and lax. That's the problem: no alarm goes off. So you sell mortgages to people who can't afford them but you're in a competitive environment, under a lot of pressure to perform, so you don't monitor the morality of your actions. And of course, as soon as that mortgage is securitized, you're not responsible for it, so you don't care if the owner defaults. Responsibility has been diffused. From your perspective, System One has worked just fine. And of course the most bonused are the most blind because for them to look carefully is, quite literally, too costly. I think when you look at how all of that plays out, you see that in conditions of high stress—from competitive environments, compensation structures, or just

159. Chugh et al., *supra* note 126, at 90.

160. See Santosh Anagol et al., *Bad Advice: Explaining the Persistence of Whole Life Insurance* 3, 16 (Feb. 8, 2011) (unpublished working paper), available at <http://ssrn.com/abstract=1786624> (finding that insurance agents in India tend to offer products that provide them high commissions, especially when the customers are less sophisticated); see also Daniel Bergstresser et al., *Assessing the Costs and Benefits of Brokers in the Mutual Fund Industry*, 22 REV. FIN. STUD. 4129, 4154–55 (2009) (studying brokers of mutual funds who supposedly help buyers select funds in which to invest and finding no tangible benefits flowing to buyers, indicating either that there are unobservable intangible benefits or (more likely) that there are material conflicts of interest between brokers and their clients).

company politics—the high stress tends to distance moral reflection.¹⁶¹

Note further that this take on bounded ethicality is completely consistent with the concept of *ethical fading*, wherein psychological forces cause ethical issues to fade from view as personal considerations come to the fore.¹⁶² Self-deception causes the ethical issue to fade from view so that an unethical decision may be unconsciously made. As Ann Tenbrunsel and David Messick explain:

This practice is common, normal, and accepted as constant and pervasive in individuals' lives. We are creative narrators of stories that tend to allow us to do what we want and that justify what we have done. We believe our own stories and thus believe that we are objective about ourselves.¹⁶³

Ethical fading is just one category of willful blindness, but an important one. Margaret Heffernan has written an entire book on the topic of willful blindness with illustrations from ethical decision making as well as every day decision making, romantic decision making, and on and on.¹⁶⁴ Consistent with previous discussions, the desire to please authority,¹⁶⁵ to fit in,¹⁶⁶ to make money,¹⁶⁷ and to preserve one's self-image as a moral person¹⁶⁸ can all cause people to be willfully blind to what is right in front of them.

The general thrust of this new body of behavioral ethics literature is that people fool themselves. They do not wake up one morning and announce, "Today is the day I start my life of crime." Rather, they often fail to realize how their self-interest colors all aspects of their decision making and then fail to observe an ethical issue sitting right in front of them. For that reason, the simple fact that stock brokers are cognizant of conflicts of interest does not mean that they will adequately adjust their behavior to account for it. Like everyone else, they have great confidence in their ethicality and their objectivity and tend to just know that they will act in a moral fashion in situations where others might not.

161. HEFFERNAN, *supra* note 105, at 159.

162. Ann E. Tenbrunsel & David M. Messick, *Ethical Fading: The Role of Self-Deception in Unethical Behavior*, 17 SOC. JUST. RES. 223, 224-25 (2004).

163. *Id.* at 225.

164. HEFFERNAN, *supra* note 105.

165. *Id.* at 107-24.

166. *Id.* at 12-45.

167. *Id.* at 182-99.

168. *Id.* at 195; *see also* Langevoort, *supra* note 66, at 102-03.

Stock brokers, like everyone else, will tend to believe that they are using moral reasoning to conclude that they can act objectively in the face of a conflict of interest, not understanding that often times moral reasoning is no more than rationalization of a previously reached conclusion that was unconsciously shaped by self-interest.¹⁶⁹ Consistent with the classic literature in motivated reasoning,¹⁷⁰ a recent study by Eric Uhlmann and colleagues provided substantial evidence that people tend to apply their moral principles malleably in order to rationalize their preferred moral conclusions.¹⁷¹ “Rather than being moral rationalists who reason from general principle to specific judgment, it appears as if people have a ‘moral toolbox’ available to them where they selectively draw upon arguments that help them build support for their moral intuitions.”¹⁷²

Needless to say, one of the strongest factors in motivated reasoning is economic influence. For example, the more people consume of a particular product (animal products, petroleum products, etc.), the more favorably they will tend to view use of that product from a moral perspective.¹⁷³ Analogously, the more stock brokers stand to earn from recommending a particular stock to customers, the more pressure they will unconsciously feel to reach a conclusion that it really is the best stock for the customer. Self-interest unconsciously drives their reasoning process.

169. See generally Haidt, *supra* note 158, at 814 (arguing that many moral decisions are made on an emotional basis and subsequent “moral reasoning” is simply an attempt to create plausible rationalizations for previously-made decisions); Eric Luis Uhlmann, et al., *The Motivated Use of Moral Principles*, 4 JUDGMENT & DECISION MAKING 479, 489 (2009) (finding substantial support for Haidt’s social intuitionist model); Piercarlo Valdesolo & David DeSteno, *Manipulations of Emotional Context Shape Moral Judgment*, 17 PSYCHOL. SCI. 476, 477 (2006) (demonstrating experimentally that by manipulating peoples’ affective states, one may also manipulate their moral judgments). On the other hand, it would probably be a significant overstatement to claim that all purported moral reasoning is nothing more than rank rationalization. See Hugo Mercier, *What Good Is Moral Reasoning*, 10 MIND & SOCIETY (forthcoming Dec. 2011).

170. See Ziva Kunda, *The Case for Motivated Reasoning*, 108 PSYCH. BULL. 480 (1990) (exploring how motivation affects reasoning); Ziva Kunda, *Motivated Inference: Self-Serving Generation and Evaluation of Causal Theories*, 53 J. PERSONALITY & SOC. PSYCHOL. 636, 637 (1987) (similar).

171. Uhlmann, et al., *supra* note 169, at 489. Uhlmann reports the results of five studies illustrating how political liberals and conservatives adjust their moral principles to reach conclusions consistent with their political biases; in other words “motivation can influence not only our descriptive beliefs about how the world is . . . but also prescriptive beliefs about how the world *ought* to be.”

172. *Id.*

173. See Robert Östling, *Economic Influences on Moral Values*, 9 B.E. J. ECON. ANALYSIS & POL’Y, issue 1, art. 8, 2009 at 1, 1 (relying in part on the operation of cognitive dissonance to explain human behavior and beliefs).

B. Moral Equilibrium

Recognizing the tendency of self-deception to enable unethical conduct to occur is an important first step in properly remedying conflicts of interest.¹⁷⁴ Unfortunately, requiring stock brokers to disclose their conflicts of interest does not help them act more ethically, as we shall see in this Section, nor does it help customers protect themselves more fully as we have already seen in the earlier discussion of insensitivity to source of information, confirmation bias, and anchoring,¹⁷⁵ and as we will see in even more striking terms in the next Section.

Some of the most interesting new studies in modern behavioral ethics deal with the concept of moral equilibrium (or moral self-regulation). The basic notion is that people constantly evaluate their own moral character and compare it to the actions they have taken. If they take an action that they view as inconsistent with their self-perceived high moral character, they often then try to “rebalance” their moral scales by subsequently acting more ethically. This is often called moral compensation.¹⁷⁶ On the other hand, if they take an ethical action that validates their self-image as a good person, they may well give themselves permission to play fast and loose with the rules for a while. This is often called moral licensing.¹⁷⁷ People actually (and unconsciously) give themselves license to depart from their own ethical standards.¹⁷⁸ In both moral compensation and moral license, “past

174. Ann E. Tenbrunsel, *Commentary: Bounded Ethicality and Conflicts of Interest*, in CONFLICTS OF INTEREST, *supra* note 97, at 96, 102.

175. See *supra* notes 70–77, 111–18, and accompanying text.

176. Evidence also shows that corporations that engage in corporate social irresponsibility are the firms most likely to then engage in corporate social responsibility. Matthew J. Kotchen & Jon Jungbien Moon, *Corporate Social Responsibility for Irresponsibility* 16 (Nat'l Bureau of Econ. Research, Working Paper No. 17254, 2011), available at <http://ssrn.com/abstract=1898510> (“We find general support for the causal relationship: when companies do more harm, they also do more good.”). The motives behind this form of moral compensation are likely more conscious and calculated than those behind individual moral compensation.

177. With moral licensing, “individuals who have validated or exceeded their ideal moral selves may experience a respite from moral regulatory forces and take ethical liberties in subsequent situations.” Chen-Bo Zhong et al., *Moral Self-Regulation*, in PSYCHOLOGICAL PERSPECTIVES ON ETHICAL BEHAVIOR AND DECISION MAKING 75, 78–79 (David De Cremer ed., 2009).

178. See generally MAX H. BAZERMAN & ANN E. TENBRUNSEL, BLIND SPOTS: WHY WE FAIL TO DO WHAT'S RIGHT AND WHAT TO DO ABOUT IT 114 (2011). Studies show that *moral hypocrisy*—manipulation of “the data so as to avoid confronting the discrepancy between one’s self-serving behavior and one’s moral standards”—is a hardy phenomenon that is consistent with the notion of moral licensing. C. Daniel Batson et al., *Moral Hypocrisy: Appearing Moral to Oneself without Being So*, 77 J.

moral choice can affect current choices, which then influence future choices."¹⁷⁹

Consider that guilt is an important moral emotion,¹⁸⁰ and that many people feel morally guilty when they purchase luxury goods that are difficult to justify.¹⁸¹ In a series of experiments Uzma Khan and Ravi Dhar gave a group of subjects (the license condition) an opportunity to commit to do a virtuous act, such as teach in a homeless shelter, improve the environment, or help a foreign student. Other subjects (the control condition) were not given these opportunities to volunteer.¹⁸² The subjects in the licensing group consistently rated themselves more positively, which is not surprising given their altruistic commitment.¹⁸³ But the result of this spike in moral self-image consistently supported the authors' hypothesis that an increase in moral self-image would lead subjects to grant themselves license to act less morally. Across a range of conditions, subjects in the license condition more frequently chose luxury goods over necessities and donated less to charity when given the opportunity (as compared to subjects in the control group).¹⁸⁴ When the volunteering was a requirement (for example, to cancel a traffic citation), the increase in self-image did not occur.¹⁸⁵ Importantly, the subjects did not consciously connect their later decisions to their earlier choice to volunteer, illustrating again the point that people tend to be unaware of how their earlier decisions (to act morally) affect their later decisions (to give themselves moral license).¹⁸⁶

PERSONALITY & SOC. PSYCHOL. 525, 527 (1999). See generally C. Daniel Batson et al., *In a Very Different Voice: Unmasking Moral Hypocrisy*, 72 J. PERSONALITY & SOC. PSYCHOL. 1335 (1997); C. Daniel Batson et al., *Moral Hypocrisy: Addressing Some Alternatives*, 83 J. PERSONALITY & SOC. PSYCHOL. 330 (2002).

179. Zhong et al., *supra* note 177, at 75.

180. JESSE J. PRINZ, *THE EMOTIONAL CONSTRUCTION OF MORALS* 18 (2007) (noting that guilt is a moral emotion that may be evolutionarily beneficial by promoting helping).

181. Darren W. Dahl et al., *The Nature of Self-Reported Guilt in Consumption Contexts*, 14 *MARKETING LETTERS* 159, 165 (2003) (finding that people often feel guilty when they buy frivolous, expensive items).

182. Uzma Khan & Ravi Dhar, *Licensing Effect in Consumer Choice*, 43 *J. MARKETING RES.* 259, 260-61 (2006).

183. *Id.* at 261.

184. *Id.* at 261-64.

185. *Id.* at 263-64.

186. *Id.* at 262. Thus, across a series of experiments, the authors found substantial evidence supporting their "main proposition . . . that engaging in or merely committing to a virtuous act can lead to an increase in a positive self-concept, which decreases the negative self-attributions associated with a luxury item and thus increases its choice likelihood." *Id.* at 260; see also Suresh Ramanathan & Patti Williams, *Immediate and Delayed Emotional Consequences of Indulgence: The Moderating Influence of Personality Type on Mixed Emotions*, 34 *J. CONSUMER RES.* 212, 220

Sonya Sachdeva and colleagues explored the question of what motivates people to act morally.¹⁸⁷ They essentially tested the notion of moral licensing and moral compensation as being related parts of an overall system of moral self-regulation.¹⁸⁸ In their first experiment, Sachdeva and colleagues asked subjects to write essays about themselves using neutral words (e.g., book, key, house), using positive words (e.g., caring, generous, fair), and using negative words (disloyal, greedy, mean). Then the three groups of subjects were given the opportunity to donate to a charity.¹⁸⁹ The subjects in the negative condition gave more than those in the neutral condition who, in turn, gave more than those in the positive condition, consistent with the notion that “participants whose moral identity was salient would donate significantly less than those whose moral identity was threatened.”¹⁹⁰ In other words, those who reminded themselves of what good people they were seemed to give themselves license not to be as charitable as they might normally have been (as indicated by the donations of the control group). On the other hand, those who were reminded of their moral failings felt the need to compensate by giving greater donations than they might normally have done. Additional experiments helped rule out causes other than moral equilibrium for the donation discrepancy¹⁹¹ and extended the findings to a setting where subjects were offered the opportunity to act cooperatively in an environmentally conscious fashion.¹⁹² Overall, the authors’ findings seem to strongly support the moral equilibrium concept.¹⁹³

(2007) (finding that consumers who indulge in a luxury good, for example, are likely to “launder” their consequent negative emotions by subsequently making utilitarian rather than hedonic choices).

187. Sonya Sachdeva, Rumen Iliev & Douglas L. Medin, *Sinning Saints and Sainly Sinners: The Paradox of Moral Self-Regulation*, 20 PSYCHOL. SCI. 523, 523 (2009).

188. *Id.* at 524 (“Our primary hypothesis was that a decrease in the moral self-concept leads to increased prosocial behavior, but that an analogous increase in the moral self-concept inhibits altruistic or prosocial behavior (relative to a neutral state).”).

189. *Id.* at 524–25.

190. *Id.* at 525.

191. *Id.* at 525–26. For example, the researchers found that when subjects were asked to write essays using positive words about *others*, their donation levels were unaffected. *Id.* at 525.

192. *Id.* at 525–26.

193. The authors concluded:

In three experiments, we found that priming people with positive and negative traits strongly affected moral behavior. We contend that these primes led participants to feel morally licensed or debased. To compensate for these departures from a normal state of being, they behaved either less morally (moral licensing) or more morally (moral cleansing).

In another interesting study, Donald Dutton and Robert Lake isolated participants who evaluated themselves as relatively unprejudiced.¹⁹⁴ They then had the participants take a test and led some of them to believe (inaccurately) that their responses had indicated the presence of racial prejudice. Other participants received feedback not indicating prejudice. As the subjects left the sessions, they were panhandled by either a black panhandler or a white panhandler.¹⁹⁵ The black panhandler received more generous donations from the subjects who had been given feedback that they were prejudiced than from those who had not been given that feedback. On the other hand, there was no material difference in donation level from the two groups to the white panhandler.¹⁹⁶ Overall, the authors found support for something comparable to a moral equilibrium effect. When given feedback that they were prejudiced, subjects who viewed themselves as unprejudiced, took advantage of an opportunity to compensate for their (imaginary) violation of their own beliefs in order to “prove to themselves that they really were not racially prejudiced.”¹⁹⁷

In an unpublished (at this writing) study, Ann Tenbrunsel and colleagues undertook a similar study and found similar results:

[S]ome participants were asked to recall items that elicited a more positive moral self-image, such as the many things they did to contribute to environmental preservation or the few things they did to contribute to environmental destruction. These participants were significantly less likely to support programs to offset carbon dioxide emissions than were participants who were asked to recall items that elicited a less positive moral self-image (i.e., the many things they did to contribute to environmental destruction or the few things they did to contribute to environmental preservation).¹⁹⁸

Id. at 527.

194. Donald G. Dutton & Robert A. Lake, *Threat of Own Prejudice and Reverse Discrimination in Interracial Situations*, 28 J. PERSONALITY & SOC. PSYCHOL. 94, 96 (1973).

195. *Id.*

196. *Id.* at 98–99.

197. *Id.* at 99.

198. BAZERMAN & TENBRUNSEL, *supra* note 178, at 115 (summarizing Ann Tenbrunsel et al., *Moral Compensation and the Environment: Affecting Individuals' Moral Intentions through How They See Themselves as Moral* (Apr. 25, 2010) (unpublished manuscript), <http://www.cognitionandculture.net/workshops/social-decisions/787-moral-compensation-and-the-environment>).

In related studies, Benoît Monin and Dale Miller first gave some subjects the opportunity to disagree with blatantly sexist statements.¹⁹⁹ When subsequently given the opportunity to make a hiring decision, these people were *more likely* to favor a man for a stereotypically male job than were subjects who had not been given the opportunity to disagree with the statements.²⁰⁰ It appeared that once individuals in the first group had established their non-sexist bona fides in their own eyes by disagreeing with the sexist statements, they gave themselves license to act in a more sexist fashion. Other iterations of the study extended the results beyond sexism to racism²⁰¹ and found that the licensing effect arose not just from the subjects' desire to appear less prejudiced to others but also to reaffirm their self-image as nonprejudiced.²⁰²

Merrill Carlsmith and Alan Gross, while studying compliance, made some subjects feel guilty by asking them to deliver apparently painful shocks to a confederate of the experimenter.²⁰³ Those subjects then appeared to feel the need to morally compensate by being particularly compliant in the experimenters' next task.²⁰⁴ Additional experiments indicated that the subjects' primary motivation was guilt rather than sympathy for their victims or a desire to make restitution.²⁰⁵

Nina Mazar and Chen-Bo Zhong found that people who bought green products gave themselves license to cheat and steal more than others who had purchased conventional products.²⁰⁶ It appeared that after buying green products and thus establishing their moral credentials, at least in their own minds, they gave themselves moral license to commit socially undesirable acts.²⁰⁷ Marketers are aware of this phenomenon—"shoppers in grocery stores usually confront the fruit-and-vegetable section first. Grocers know that shoppers who buy the healthy stuff first will feel so uplifted they will buy more junk food later in their trip."²⁰⁸

199. Benoît Monin & Dale T. Miller, *Moral Credentials and the Expression of Prejudice*, 81 J. PERSONALITY & SOC. PSYCHOL. 33, 34-35 (2001).

200. *Id.* at 35-36. The effect only manifested itself in male subjects, however. *Id.* at 35.

201. *Id.* at 36-37.

202. *Id.* at 39.

203. J. Merrill Carlsmith & Alan E. Gross, *Some Effects of Guilt on Compliance*, 11 J. PERSONALITY & SOC. PSYCHOL. 232, 233 (1969).

204. *Id.* at 235-36.

205. *Id.* at 237-38.

206. Nina Mazar & Chen-Bo Zhong, *Do Green Products Make Us Better People?*, 21 PSYCHOL. SCI. 494, 497 (2010).

207. *Id.* at 498.

208. BROOKS, *supra* note 55, at 171.

All these studies provide evidence for the phenomenon of moral self-regulation (moral equilibrium) and its two aspects, moral compensation and moral license.²⁰⁹ The obvious implications of this process of moral equilibrium for the current issue are profound. These studies indicate that if stock brokers have just done something upfront and honest (disclosed conflicts of interest), they may tend to unconsciously give themselves moral license to take a little advantage of their customers.²¹⁰ Their overall treatment of their customers will thus be made worse, not better, by the disclosure requirement. This is not a conscious decision; most people are unaware of how their ethical decision making can be influenced by a previous decision they have made.²¹¹

As it happens, recent studies have indicated that these obvious implications are very, very plausible.

1. DISCLOSURE DISTORTION

Daylian Cain, George Loewenstein, and Don Moore studied moral equilibrium in a context that has direct relevance for the issues addressed in this Article.²¹² Traditional economic theory indicates that disclosure of a conflict of interest should protect the disclosees who can in response adjust their judgments regarding the value of the advice given. Cain and colleagues, however, hypothesized that for several reasons disclosees might not know what to do with the information disclosed.²¹³ First, the dangerous bias in this setting more frequently is not the conscious attempt of the doctor, stock broker, or other discloser to rip off the disclosees; rather, as noted above, the bigger problem is with unconscious bias, and people (the disclosees) generally fail to appreciate the impact of unconscious bias and the extent to which it can distort advice that is given.²¹⁴ Second, as also noted earlier,²¹⁵ “people tend to be naturally trusting and credulous toward their own

209. There are other interesting, related studies. For example, Brock Bastian and colleagues found that subjects who wrote about a time that they behaved unethically became motivated to experience pain as a means of “rebalancing the scales.” Brock Bastian et al., *Cleansing the Soul by Hurting the Flesh: The Guilt-Reducing Effect of Pain*, 22 *PSYCHOL. SCI.* 334, 335 (2011).

210. See BAZERMAN & TENBRUNSEL, *supra* note 178, at 116 (referring to the process as “psychological license”).

211. Zhong et al., *supra* note 177, at 84.

212. See Daylian M. Cain, George Loewenstein & Don A. Moore, *The Dirt on Coming Clean: Perverse Effects of Disclosing Conflicts of Interest*, 34 *J. LEGAL STUD.* 1 (2005).

213. *Id.* at 5.

214. *Id.*

215. See discussion *supra* Part II.C-G.

advisors.”²¹⁶ Third, when people do adjust their judgments in response to disclosure of a conflict of interest they tend to do so insufficiently. Among natural human tendencies that account for this are anchoring and adjustment (when people are anchored on a particular conclusion, they tend to adjust insufficiently to introduction of new information)²¹⁷ and the “curse of knowledge” (the difficulty people have in factoring out information they already know has become irrelevant or discredited).²¹⁸

Cain, Loewenstein, and Moore also hypothesized that disclosure could *worsen* advisors’ performance in two ways. First, while having to disclose conflicts of interest might incentivize advisors to consciously try to improve their accuracy, they might also be tempted to provide even more biased advice knowing that discloseses would perhaps discount their advice somewhat due to the disclosure.²¹⁹ Second, advisors might unconsciously grant themselves the moral license mentioned above.²²⁰ Disclosing the existing conflicts of interest could reduce advisors’ guilt regarding their conflicts of interest and cause them to unconsciously give even more biased advice.²²¹

Cain and colleagues decided to test these theories empirically. They set up an experiment where “estimators” were rewarded for their accuracy in estimating an uncertain quantity.²²² These participants played the role, roughly, of the recipient of advice (broker’s customer, doctor’s patient, attorney’s client, etc.).²²³ The experiment also featured “advisors” who provided information to the estimators.²²⁴ In a control treatment, the advisors received higher rewards if the estimators’ guesses were more accurate.²²⁵ In two versions of a conflict-of-interest treatment, advisors were paid more when the estimator responded with a high, rather than an accurate, estimate.²²⁶ In both versions of this

216. Cain, Loewenstein & Moore, *supra* note 212, at 5.

217. See generally Gretchen B. Chapman & Eric J. Johnson, *Incorporating the Irrelevant: Anchors in Judgments of Belief and Value*, in *HEURISTICS AND BIASES*, *supra* note 129, at 120, 120–38.

218. See John C. Anderson et al., *Evaluation of Auditor Decisions: Hindsight Bias Effects and the Expectation Gap*, 14 *J. ECON. PSYCHOL.* 711, 715–20 (1993) (discussing one aspect of the curse of knowledge—the hindsight bias, which is the inability to discount how events turned out when judgment decisions were made beforehand).

219. Cain, Loewenstein & Moore, *supra* note 212, at 6–7.

220. *Id.* at 7.

221. *Id.*

222. *Id.* at 8.

223. *Id.*

224. *Id.*

225. *Id.*

226. *Id.*

treatment, the interests of the advisors and the estimators diverged. In one version of the conflict-of-interest treatment, advisors disclosed their conflict of interest; in the other they did not.²²⁷

The study found that advisors who disclosed their conflict of interest gave less accurate advice than did those who did not disclose their conflict of interest (and both, naturally, gave less accurate advice than did advisors whose interests were aligned with the estimators).²²⁸ Estimators who were informed of the conflict of interest discounted the advice somewhat, but not sufficiently to offset the increase in the bias of the advice they received.²²⁹ Unsurprisingly, estimators made more money where there was no conflict of interest.²³⁰ Where there was a conflict of interest, estimators made more money where the advisors' conflict of interest was *not* disclosed than when it was.²³¹ The authors' ultimate conclusion was:

Disclosure, at least in the context of the admittedly stylized experiment discussed in this paper, benefited the providers of information but not its recipients. To the extent that a similar effect occurs outside the experimental laboratory, disclosure would supplement existing benefits already skewed toward information providers. In particular, disclosure can reduce legal liability and can often forestall more substantial institutional change. We do not believe that this is a general result—that is, that disclosure always benefits providers and hurts recipients of advice—but it should challenge the belief that disclosure is a reliable and effective remedy for the problems caused by conflicts of interest.²³²

Thus, required disclosure can enable a stock broker, for example, to unconsciously rationalize selling a product to clients that is not the best available because, after all, he or she has warned the clients of their vulnerable situation.²³³ “Disclosure distortion” is not an inevitable

227. *Id.* at 7–8. The three experimental conditions were called the “accurate” condition, the “high/disclosed” condition, and the “high/undisclosed” condition. *Id.* at 10.

228. *Id.* at 17.

229. *Id.*

230. *Id.* at 18.

231. *Id.*

232. *Id.* at 20. It is worth emphasizing that the authors do not believe that the result of this experiment “is a general result.” *Id.*

233. Cain and colleagues suggest another reason for this overall result. They suggest that stock brokers who must disclose their conflict of interest will reason that customers will discount their advice somewhat. Cain et al., *supra* note 9797, at 115. Because they really want the customers to follow their advice, they may engage in

result of mandatory NFD disclosure, but it is obviously a matter of concern.²³⁴ This is especially so because Christopher Robertson has replicated Cain and colleagues' findings in a separate experiment, finding virtually identical results,²³⁵ although some other studies found some mitigating factors.²³⁶

Robertson's experiment tried many versions of disclosure to try to improve results. It studied early disclosure versus late disclosure, disclosure attempting to inform the customer regarding the degree of existing bias, etc.²³⁷ After all that, Robertson still found that "[a]s in [the Cain et al. study], a mandatory disclosure policy does not seem to help laypersons adjust their reliance on the advice received. Instead, it may only cause the expert advisors to become more biased."²³⁸

Cain, Loewenstein, and Moore themselves performed another round of more sophisticated studies²³⁹ and again confirmed earlier findings "that disclosure lessens moral reluctance to provid[e] biased advice,"²⁴⁰ that "disclosure [leads] to greater distortion of advice when there [is] a conflict of interest,"²⁴¹ that the small amount of discounting of advice that recipients of the disclosure undertake is "insufficient to compensate for the increased bias" in the advice,²⁴² and that, the bottom line is that "disclosure hurt the financial outcomes of those it was supposed to protect."²⁴³

"strategic exaggeration." *Id.* The brokers are used to saying eighty and having the customers accept eighty. *See id.* Because the required NFD disclosure may cause customers to discount the broker's advice, the broker may well say a hundred to induce the customers to accept eighty. Thus, the required disclosure may lead to exaggerated advice. *Id.*

234. *Id.* at 117 ("Rather than show that disclosures always exacerbate the problems created by conflicts of interest, our goal has been to argue that disclosure cannot be assumed to always help.")

235. Robertson, *supra* note 28, at 669 ("All of the point estimates in the present study are statistically indistinguishable from those in the [Cain et al.] study.")

236. *See* Bryan Church & Xi (Jason) Kuang, *Conflicts of Interest, Disclosure, and (Costly) Sanctions: Experimental Evidence*, 38 J. LEGAL STUD. 505, 527 (2009) (finding that disclosure in combination with sanctions can reduce adverse effects); Christopher Koch & Carsten Schmidt, *Disclosing Conflicts of Interest—Do Experience and Reputation Matter?*, 35 ACCT., ORGS. & SOC'Y 95, 101-05 (2010) (finding evidence of moral licensing in auditor setting, but also finding effect may diminish with experience and reputation).

237. Robertson, *supra*, note 28, at 678, 702 tbl.3.

238. *Id.* at 670.

239. Daylian M. Cain, George Loewenstein & Don A. Moore, *When Sunlight Fails to Disinfect: Understanding the Perverse Effects of Disclosing Conflicts of Interest*, 37 J. CONSUMER RES. 836, 837 (2011).

240. *Id.* at 842.

241. *Id.* at 844.

242. *Id.*

243. *Id.* at 845.

2. DISCLOSURE, THE PANHANDLER EFFECT, AND INSINUATION ANXIETY

An additional study by Cain and Loewenstein with Sunita Sah rendered these findings even more vexing. They tested notions including the “panhandler effect” and “insinuation anxiety.”²⁴⁴

Consider that if a little neighborhood girl comes to your door to sell you some candy, you are probably less likely to buy if she says “Would you like some candy?” than if she says: “Would you like some candy? If you buy I’ll make my sales target and earn a trip to Disneyland.” When people learn of others’ interests, they often buy (or give) because they feel added pressure to do so and do not wish to appear unfair or ungenerous.²⁴⁵ That is the *panhandler effect*.

As a matter of social etiquette, it is extremely rude to question another’s integrity.²⁴⁶ Few actions are more impolite or evoke stronger emotional reactions.²⁴⁷ It is so against the nature of most people to question others’ integrity that they will often choose not to do so even in the face of blatant evidence of dishonesty. In *Kravitz v. Pressman, Frohlich & Frost, Inc.*,²⁴⁸ a stock broker was rather egregiously churning his client’s account and slowly reducing her balance to near zero.²⁴⁹ But whenever she questioned him or asked for information, he would become angry and claim that he was acting in her best interests, and she would back off.²⁵⁰ People are leery of questioning others’ integrity, so when they are informed by advisors of the advisors’ conflict of interest, they may be *more* likely to accept the advice, because to refuse it signals distrust and thereby insults the discloser.²⁵¹ This is *insinuation anxiety* in action.

244. Sunita Sah, George Loewenstein & Daylian M. Cain, *The Burden of Disclosure: Increased Compliance with Distrusted Advice* 6–7 (Mar. 12, 2011) (unpublished manuscript) (on file with the *Wisconsin Law Review*).

245. *Id.* at 8–11.

246. See Larry J. Cohen & Joyce H. Vesper, *Forensic Stress Disorder*, 25 *LAW & PSYCHOL. REV.* 1, 10 (2001) (observing that questions that challenge one’s honesty are considered insulting).

247. See JOSHUA FOER, *MOONWALKING WITH EINSTEIN: THE ART AND SCIENCE OF REMEMBERING EVERYTHING* 231–32 (2011) (stating that when he accused purported savant Daniel Corney of being a fraud, “[i]t was one of the most uncomfortable sentences I’d ever spoken to anyone”); Ewa Bogdanowska-Jakubowska, *Cultural Variability in Face Interpretation and Management*, in *POLITENESS ACROSS CULTURES* 237, 247–48 (Francesca Bargiela-Chiappini & Dániel Z. Kádár eds., 2011) (giving examples in English and Polish cultures).

248. 447 F. Supp. 203 (D. Mass. 1978).

249. *Id.* at 207–09.

250. *Id.* at 208.

251. See Sah, Loewenstein & Cain, *supra* note 244, at 6–7.

Sah, Loewenstein, and Cain addressed both the panhandler effect and insinuation anxiety as falling within the broader notion of a *burden-of-disclosure* effect: "Disclosure places advisees in an effective bind because it causes advisees to trust the advice less but also experience increased pressure to comply with the advice."²⁵²

In a series of experiments, Sah, Loewenstein, and Cain found that when a conflict of interest was disclosed by an advisor, the advisee *tended to trust the advice given less, but to accept it more.*²⁵³ In the experiment, "choosers" were given a choice of two options. A was clearly superior to B.²⁵⁴ When the advisors' and the choosers' interests were aligned, choosers selected the superior A option 93% of the time.²⁵⁵ However, when the advisors' interests were served by the choosers selecting option B and they advised the choosers to do so, the choosers then chose the inferior option not 7% of the time, but 52% of the time.²⁵⁶ The real point of the experiment, however, was that when advisors not only advised the choosers to choose the inferior option but disclosed that they (the advisors) would receive a greater reward if the choosers chose the inferior option, the choosers chose the inferior option 81% of the time.²⁵⁷

The researchers had several iterations of the experiment to tease out causal factors, concluding:

The results of our six experiments show that people experience conflicting emotions when receiving disclosure of a [conflict of interest] from an advisor. . . . Advisees are simultaneously aware that the advice is likely to be biased and trust it less, yet feel increased pressure to comply with the advice. Hence, instead of a warning, disclosure can become a burdensome request to comply with advice that is trusted less.

Two mechanisms, insinuation anxiety and the panhandler effect, contribute to this increased pressure to comply, although the magnitude of each of these mechanisms will inevitably vary by context.²⁵⁸

252. *Id.* at 3.

253. *Id.*

254. *Id.* at 20.

255. *Id.* at 22.

256. *Id.* at 23.

257. *Id.*

258. *Id.* at 40-41.

This turns out to be yet another in a series of studies indicating that mandatory disclosure is not adequate to the task in the stock broker setting. Indeed, it may well be counterproductive.²⁵⁹

CONCLUSION

Lawmakers and regulators seem to default to disclosure as the primary remedy for even the most dramatic ills of the financial and securities systems. This is a natural thing for them to do because disclosure appears on its face to be cheap, easy, and effective.²⁶⁰ “In short, when lawmakers are besieged, mandated disclosure looks like rescue. Its critics are few. Lawmakers can be seen to have acted. The fisc is unmolested.”²⁶¹

Unfortunately, disclosure mandated by government often does not do the job with which it is charged.²⁶² There are now many, many studies which tend to indicate that mandated disclosure as a remedy—whether one addresses physicians’ disclosures to patients, lenders’ disclosures to borrowers, broker-dealers’ disclosure to investors, or cops’ disclosures to criminals (the Miranda warnings)—is often ineffective. Ben-Shahar and Schneider surveyed scores of empirical studies and reached the discouraging conclusion that:

Although mandated disclosure addresses a real problem and rests on a plausible assumption, it chronically fails to accomplish its purpose. Even where it seems to succeed, its costs in money, effort, and time generally swamp its benefits. And mandated disclosure has unintended and undesirable consequences, like driving out better regulation and hurting the people it purports to help.²⁶³

This conclusion derives additional support from the recent global financial crisis which has highlighted the limits of a disclosure regime

259. See Cain, Loewenstein & Moore, *supra* note 239, at 848–49 (“Disclosure can also change the nature of interactions between parties, making it more uncomfortable for an advice recipient to reject advice.”).

260. Ben-Shahar & Schneider, *supra* note 47, at 681–82.

261. *Id.* at 684; see also FUNG ET AL., *supra* note 21, at 15 (noting that sometimes disclosure rules “create an illusion that a problem has been addressed while producing minimal impact”).

262. See Ripken, *supra* note 17, at 189 (“Congress and the SEC always seem to tackle thorny problems of corporate governance with some variant of disclosure rather than face the merits of these problems head-on.”).

263. Ben-Shahar & Schneider, *supra* note 47, at 651.

in that in many areas where tragedy struck proper information was available for all to see, yet the disclosure did not prevent disaster.²⁶⁴

Therefore, it should not be a surprise that a law that says: "Don't subordinate your customers' interests to your own" will have more beneficial impact than a law that says: "Tell customers that you do not have to subordinate your interests to theirs." The excellent popular writer James Surowiecki said it best in the following oft-quoted passage:

It has become a truism on Wall Street that conflicts of interest are unavoidable. In fact, most of them only seem so, because avoiding them makes it harder to get rich. That's why full disclosure is suddenly so popular: it requires no substantive change. . . . Transparency is well and good, but accuracy and objectivity are even better. Wall Street doesn't have to keep confessing its sins. It just has to stop committing them.²⁶⁵

Without examining all of the issues raised by the proposal to consider a uniform fiduciary duty for stock brokers and other financial professionals, this Article agrees with Surowiecki that disclosure alone seems a frail tool with which to attack the many ills that arise from blatant conflicts of interest in the financial industry. The multitude of

264. Avgouleas, *supra* note 22, at 14–22. Avgouleas makes a strong case for an explanation for the global financial crisis that is unrelated to disclosure deficiencies, noting that:

[I]nvestors had in many cases sufficient information about the risks of their investment strategies and of the financial products used to implement them. Yet market actors could not properly process available information in those cases and adjust their positions to the riskiness of structured credit securities for a variety of reasons. First, due to product complexity, boundedly rational investors failed to understand the mechanics and risks of shadow banking and structured credit securities. Second, because of market players' tendency to herd, responding strategically to other market actors' behaviour, these did not have the capacity or the desire to use in a rational way the disclosed information and take contrarian positions. Third, the influence of other behavioural factors such as the use of heuristics, and investor overconfidence in times of market euphoria, because of abundance of easy credit and rising market prices, meant that investors chose to ignore the warning signals in the disclosed data in favour of over-reliance on credit ratings.

Id. at 4 (footnotes omitted).

265. James Surowiecki, *The Talking Cure*, NEW YORKER, Dec. 9, 2002, at 54; see also Cain et al., *supra* note 97, at 107 ("Part of the reason [disclosure seems such an attractive policy] is that firms and individuals facing conflicts of interest often view it as the lesser of evils: they would rather have to tell about their conflicts of interest than have to get rid of them through divestiture or recusal.").

studies surveyed in this Article may not be definitive, but they certainly provide reason to doubt the efficacy of the disclosure remedy in this setting.

No-fiduciary-duty disclosures do not provide brokers' customers sufficient material with which to adequately assess their situation and protect themselves from potential abuse. Nor do NFD disclosures adequately rein in brokers. Rather, they may well cause brokers to unconsciously give customers even more biased advice than they would have given in the absence of the disclosures, as they unconsciously grant themselves moral license to depart from their own ethical standards.²⁶⁶ Indeed, the disclosures, through the panhandler effect and insinuation anxiety, may well cause customers to be even more likely to follow biased advice than they would have had the disclosures not been made. In short, the mandated NFD disclosures, several empirical studies show, are likely to worsen the effects of the inherent conflict-of-interest situation.

266. *See generally* WRAY HERBERT, ON SECOND THOUGHT: OUTSMARTING YOUR MIND'S HARD-WIRED HABITS 242 (2010) ("So it appears that our inner moralist deals in a kind of moral 'currency.' We collect chits through our good deeds and debts through our transgressions, and we spend our chits to pay off our moral debts. That way, we keep the moral ledger balanced.").

Copyright of Wisconsin Law Review is the property of Wisconsin Law Review and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.

86

BIASED ADVICE

*Christopher Tarver Robertson**

ABSTRACT

The modern capitalist society, characterized by decentralized decision making and increasingly sophisticated products and services, turns on relationships of epistemic reliance, where laypersons depend upon advisors to guide their most important decisions. Yet many of those advisors lack real expertise and many are biased by conflicting interests. In such situations, laypersons are likely to make suboptimal decisions that sometimes aggregate into systematic failures, from soaring health care costs to market crashes. Regulators can attempt to manage the symptoms and worst abuses, but the fundamental problem of biased advice will remain. There are many potential policy solutions to the fundamental problem, from outright bans on conflicting interests to disclosure mandates, yet their comparative effectiveness is poorly understood.

By constructing a decision task for human subjects and providing advice in various scenarios, this Article reports new field experiments testing alternative policy mechanisms. Prior research has shown that disclosure mandates can be deleterious if they make advisors more biased, but this paper contextualizes those findings. It turns out that disclosures may be valuable in settings where relative expertise is low, but deleterious where relative expertise is high. By also disaggregating the data, one finds that disclosures of conflicting interests may hurt laypersons in the majority of situations where the conflicted advice is not actually biased. Thus, the evidence suggests that, if they are to be at all effective, disclosure mandates should be narrowly tailored.

Most importantly, the evidence shows that a disclosure mandate improves layperson performance when unbiased advisors are also available. Yet

* Associate Professor, James E. Rogers College of Law, University of Arizona: chris.robertson@law.arizona.edu. This study was funded by the Petrie Flom Center at Harvard Law School. The author thanks Daylian Cain, Tess Wilkinson-Ryan, Ken Carson, Melissa Wasserman, Brian Sheppard, Jamie Robertson, I. Glenn Cohen, Susannah Rose, Michael Frakes, Vincent Chiao, Kathie Barnes, Brent White, the participants in the Harvard Medical School Bioethics Works in Progress Series, the Conference on Empirical Legal Studies, and the University of Arizona James E. Rogers College of Law Works in Progress Series. Nicholas Perros, Geoff Balon, and David Yokum provided excellent research assistance.

laypersons appear to be poor judges of their need for unbiased advice, so market mechanisms may be ineffective for provisioning unbiased advice. In the end, the presence of an unbiased advisor is the strongest determinant of layperson performance, and thus policymakers must develop ways of aligning the interests of advisors and laypersons. Pay-for-performance, blinding of experts, and mandatory or subsidized second-opinion policies are likely to be helpful in aligning these interests.

INTRODUCTION	655
A. <i>The Problem</i>	655
B. <i>Potential Policy Solutions</i>	659
I. HOW A MANDATORY DISCLOSURE POLICY CAN HURT LAYPERSONS BY DEGRADING THE ADVICE GIVEN	666
A. <i>The Cain, Loewenstein, and Moore Study (CLM)</i>	666
B. <i>The Present Experiment's Replication and Extension of CLM</i> ..	668
II. WHEN A DISCLOSURE, OR EVEN A BAN, MIGHT WORK, DEPENDING ON RELATIVE EXPERTISE AND DEGREE OF BIAS	670
A. <i>Measuring Epistemic Asymmetry and Bias</i>	670
B. <i>Extrapolating to Real World Conditions to Test Policy Solutions</i>	673
III. MAKING DISCLOSURES WORK BETTER THROUGH ANCHORING, INFORMATION TECHNOLOGY, AND PERSONALIZATION	677
A. <i>When to Disclose</i>	678
B. <i>What to Disclose</i>	679
C. <i>To Whom to Disclose</i>	683
IV. CALIBRATING RELIANCE IN A MARKET FOR ADVICE	686
A. <i>Affirmative Disclosures of Aligned Interests</i>	686
B. <i>Using Disclosures to Select Advisors</i>	688
C. <i>The Value of Second Opinions</i>	690
D. <i>A Market for Unbiased Advice</i>	691
CONCLUSIONS—ELIMINATING BIASES WITH SOUND POLICY	696
METHODOLOGICAL APPENDIX	698

INTRODUCTION

A. The Problem

Atul Gawande recently illustrated the economics of the practice of medicine in America by profiling one area—McAllen, Texas—which leads the nation in the problem of increasing health care costs without observable increases in quality:

General surgeons are often asked to see patients with pain from gallstones. If there aren't any complications—and there usually aren't—the pain goes away on its own or with pain medication.

... A surgeon has to provide reassurance (people are often scared and want to go straight to surgery), some education about gallstone disease and diet, perhaps a prescription for pain; in a few weeks, the surgeon might follow up. But increasingly, I was told, McAllen surgeons simply operate. The patient wasn't going to moderate her diet, they tell themselves. The pain was just going to come back. And by operating they happen to make an extra seven hundred dollars.¹

This vignette depicts a situation of epistemic reliance.² The surgeon has a much better ability to determine how best to treat gallbladder pain compared to the patient, a layperson untrained in medicine, and the patient thus reasonably relies upon the surgeon for advice. This vignette also depicts conflicting interests, where the surgeon is in part motivated (perhaps only subconsciously) by the prospect of receiving payment for the service of surgery, while the patient instead seeks health and, all other things being equal, prefers to avoid the expenses, pain, inconvenience, and risks of needless surgery. Whether these conflicting interests cause surgeons to make different recommendations than they would otherwise make, i.e., whether the conflicts cause biases, is an empirical question.³

¹ Atul Gawande, *The Cost Conundrum*, NEW YORKER, June 1, 2009, at 36, 36, 38.

² Epistemology is the philosophical study of knowledge, i.e., how persons develop justified true beliefs. I call the expert–layperson relationship “epistemic reliance” because the layperson is unable to directly assess the truth, but instead must rely upon the advisor who is more able to do so. See generally *THE PHILOSOPHY OF EXPERTISE* (Evan Selinger & Robert P. Crease eds., 2006) (collecting essays exploring this epistemic relationship).

³ See Alan L. Hillman et al., *How Do Financial Incentives Affect Physicians' Clinical Decisions and the Financial Performance of Health Maintenance Organizations?*, 321 NEW ENG. J. MED. 86, 86 (1989) (reviewing the evidence).

These sorts of situations, where informational asymmetry exists between doctor and patient, and their motivations are out of sync, can be found throughout medicine. As one recent report explained the general problem:

[C]onsumers . . . face a huge knowledge gap compared with care providers and are therefore highly reliant—and understandably so—on the advice and guidance of their physicians. In the absence of evidence to the contrary, patients may often assume that more care, or more expensive care, will lead to better outcomes.

. . . [Meanwhile, fee-for-service reimbursement, the primary method of payment for outpatient care, . . . creates financial incentives [for physicians] to provide more care, and care that is more costly. More visits, more tests, more procedures all add up to more pay for providers and higher costs to the system.⁴

In the aggregate, as laypersons' choices are systematically skewed by such biased advice, the problem creates massive externalities and systematic failures. While serving as the director of the Congressional Budget Office, Peter Orszag argued that "our country's financial health will in fact be determined primarily by the growth rate of per capita health care costs," and he pointed at fee-for-service incentives as a primary cause.⁵ The health care industry is characterized by radically distributed decision making, with each patient deciding upon her own course of treatment within the range of treatments offered by providers and covered by public and private insurers. Thus, real reform of health care costs may need to focus on fixing the relationship of epistemic reliance and the conflicting interests at the bottom levels of the health care economy, since that is where the decisions are made.

For another example of this problem of epistemic reliance and bias, consider the wave of home mortgage foreclosures that contributed to the "Great Recession." In the wake of the mortgage-lending debacle, which rocked global financial markets and caused policymakers to make unprecedented interventions in the financial industry, the Federal Deposit Insurance Corporation took a hard look at the subprime lending products and practices of the mortgage industry.⁶ Were too many loans being made to unqualified borrowers? Were the exotic mortgage products destined to fail?

⁴ DIANA FARRELL ET AL., MCKINSEY GLOBAL INST., ACCOUNTING FOR THE COST OF U.S. HEALTH CARE 28, 31 (2008), available at http://www.mckinsey.com/mgi/reports/pdfs/healthcare/US_healthcare_report.pdf.

⁵ Peter R. Orszag & Philip Ellis, *The Challenge of Rising Health Care Costs—A View from the Congressional Budget Office*, 357 NEW ENG. J. MED. 1793, 1793–94 (2007).

⁶ See generally Ryan Lizza, *The Contrarian*, NEW YORKER, July 6, 2009, at 30 (describing governmental responses to the foreclosure crisis).

The financial industry executives demurred about their practices and products, pointing instead toward the decentralized decisions made by every homebuyer taking out a mortgage and every homeowner considering a refinance.⁷ The executives said, “You know, it’s kind of like the N.R.A.—people kill people, not guns! It’s not the mortgages, it’s the borrowers.”⁸

There is some truth in that demurrer. Notwithstanding all the regulations at the margins, a mortgage agreement is ultimately a contract, founded on the idea of voluntarily chosen promises.⁹ Borrowers can bind themselves for decades to whatever financial products the banks want to offer them, and if the borrowers make bad decisions, then they suffer the consequences, along with the banks that made the bad bets when they issued the mortgages to those borrowers.

The borrower-centric analysis ignores the reality of epistemic reliance and conflicting interests, which underlie these transactions. Borrowers have little ability to interpret voluminous and technical loan documents. Nor can they compare the real costs of the various contractual terms or use actuarial data to weigh the likelihood of defaulting, given various economic scenarios over the next few decades. As Elizabeth Warren explains,

The effective deregulation of interest rates, coupled with innovations in credit charges (e.g., teaser rates, negative amortization, increased use of fees, cross-default clauses, penalty interest rates, and two-cycle billing), have turned ordinary credit transactions into devilishly complex financial undertakings. Aggressive marketing, almost nonexistent in the 1970s, compounds the difficulty, shaping consumer demand in unexpected and costly directions. And yet consumer capacity—measured both by available time and expertise—has not expanded to meet the demands of a changing credit marketplace.¹⁰

As a result, borrowers can either fly blind or rely upon the advice of others, most frequently mortgage brokers, who purportedly have expertise, experience, and information about the mortgage market, which the borrowers lack.¹¹

⁷ *See id.*

⁸ *Id.* at 34 (internal quotation marks omitted).

⁹ *See* FDIC v. Hennessee, 966 F.2d 534, 537 (10th Cir. 1992) (“[A] mortgage is a contract and is generally subject to the rules of construction applicable to contracts.”).

¹⁰ Elizabeth Warren, *Unsafe at Any Rate*, DEMOCRACY, Summer 2007, at 8, 10.

¹¹ *See id.* at 12 (noting mortgage brokers’ advertisements, e.g., “a friend to help you find the best possible mortgage” (internal quotation marks omitted)).

The borrower–broker reliance relationship is, however, skewed by conflicting interests. As Joseph Stiglitz explains in his postmortem on the causes of the Great Recession, mortgage brokers “were *supposed* to be working for the borrower, but they often received kickbacks from the lender—an obvious conflict of interest. . . . Worse, the brokers got the biggest rewards for steering borrowers into the riskiest mortgages, adjustable-rate loans with prepayment penalties, and even got kickbacks when the borrower refinanced.”¹² In short, their advice was biased. If too many of these mortgages are being issued to unqualified borrowers, or if too many of these mortgages are defaulting, the brokers may be a major cause.¹³ One scholar explains that this

[i]nformation asymmetry [between borrower and broker] enables a predatory lender or mortgage broker to exert dominance over the borrower in the initial marketing of the loan and to insert into the loan documents terms that produce destructive effects, such as stripping the borrower’s equity in her property or creating conditions that too often make foreclosure inevitable.¹⁴

If this account is correct, it expands the policymaker’s inquiry beyond the legalistic notion of a contract as the voluntary promises of two parties and instead demands attention to the epistemic context in which these decisions are made, a context centered on a biased advisor.¹⁵ As long as the primary decisionmakers in this economic system lack the epistemic resources to make wise decisions by themselves, and as long as their advisors are motivated by interests other than the well-being of the decisionmakers, it seems that individual failures and systematic problems are inevitable.

Escalating health care costs and the crashing mortgage finance sector are just two of the most obvious examples of the problem of biased expertise,

¹² JOSEPH E. STIGLITZ, *FREEFALL* 89 (2010); see also Michael S. Barr et al., *Behaviorally Informed Home Mortgage Credit Regulation*, in *BORROWING TO LIVE* 170, 175–76 (Nicolas P. Retsinas & Eric S. Belsky eds., 2008) (describing incentives for mortgage brokers to steer reliant borrowers to more expensive loan options); Warren, *supra* note 10, at 12–13 (describing the type of “broker who is working only for himself, taking what amounts to a bribe from a mortgage company to steer a family into a higher-priced mortgage than it could qualify for, all the while assuring the family that this is the best possible deal”).

¹³ See Lloyd T. Wilson, Jr., *Effecting Responsibility in the Mortgage Broker-Borrower Relationship: A Role for Agency Principles in Predatory Lending Regulation*, 73 U. CIN. L. REV. 1471 (2005).

¹⁴ *Id.* at 1473 (footnotes omitted).

¹⁵ See generally Gillian K. Hadfield et al., *Information-Based Principles for Rethinking Consumer Protection Policy*, 21 J. CONSUMER POL’Y 131, 140 (1998) (“Perhaps the most important lesson that emerges from modern bargaining theory is the essential role that information, and in particular information asymmetry, plays in bargaining.”).

where laypersons relied upon advisors to make some of the most important decisions in their own lives, but received bad advice that aggregated into systemic failures. Without looking beyond the front pages of the daily newspaper, one can find many other examples of this problem.¹⁶ Indeed, one might argue that these sorts of decisions are archetypical of modern capitalism, which is defined by distributed, decentralized decision making. It depends on each farmer, each household, each worker, and each business to make their own more or less rational decisions as to their own consumption and production functions. As society becomes increasingly complex—as new medical treatments are discovered and new financial instruments are crafted, as new chemicals are put into our foods and as new high-tech tools are deployed in our workplaces—distributed decisionmakers must rely upon specialists who have developed expertise in understanding and using these sophisticated products. The economics of those advisory relationships then become the central questions for understanding the economics of society.

B. Potential Policy Solutions

When these ground-level problems between laypersons and their biased advisors bubble up into system-wide crises, policymakers may search for solutions. A reflexive answer is to implement top-down substantive regulations of affected industries. Regulators will and should aim for seemingly low-hanging (but rotten) fruit that can be easily lopped-off—i.e., banning those products that are little more than “tricks and traps” for consumers and that “have no place in a well-functioning market.”¹⁷ These are the products whose costs or risks are so obviously out of proportion to the benefits that no well-informed consumer would ever utilize them. Whether it

¹⁶ For another example of this dynamic, scholars of the accounting industry explain:

Conflicts of interest played a central role in the corporate scandals that shook America at the turn of the twenty-first century. Many companies have joined Enron and WorldCom in issuing earnings restatements as a result of inaccuracies in published financial reports. . . . At the root of both this mismanagement and the failure of monitoring systems lie conflicts of interest. . . . Accounting firms have incentives to avoid providing negative audit opinions to the managers who hire them and pay their auditing fees.

Don A. Moore et al., *Conflicts of Interest and the Case of Auditor Independence: Moral Seduction and Strategic Issue Cycling*, 31 ACAD. MGMT. REV. 10, 10 (2006). Investors rely upon these accounting firms' privileged access and special expertise in evaluating company finances, and simply must do so. Yet, the reliance relationship is undermined by such predictable biases. Similarly, scholars have pointed to the systematic biases that realtors insert into the real estate market, which were likely responsible for exacerbating the real estate bubble and also helped to destroy billions of dollars of net worth held by individual citizens.

¹⁷ Warren, *supra* note 10, at 11.

is an onerous term in a mortgage document or an unproven drug, the government is sometimes willing to substitute its judgment for those of the consumer and simply ban that transaction.¹⁸ Let us call this general category of regulations, which specifically focus on the appropriateness of products and services, “substantive” regulations.¹⁹

Substantive regulation has limits. First, this sort of governmental paternalism is anathema to deeply held values. In medicine most clearly, it has long been understood that “[e]very human being of adult years and sound mind has a right to determine what shall be done with his own body.”²⁰ There are also epistemic problems. Many products and services may be good for some consumers in some situations, but bad for others in other situations, which makes it quite difficult for the regulator to effectively control the substance of the transaction by *ex ante* decree. A given treatment may only work for 10% of patients, but the difficult question for the surgeon and the layperson is to determine whether *this patient* will be in the 10% or the 90%.²¹ If laypersons could simply follow a rote guideline to decide whether to undergo surgery, or to determine which mortgage to buy, they would not need the expert’s advice at all. Thus, the very category of cases where biased advice is the problem is also the category of cases where substantive regulation is least likely to be effective. In these contexts, substantive regulation becomes a blunt instrument, doing harm as often as it does good.²²

Substantive regulation also faces a moving target. By simply capping interest rates, regulators of the consumer financial sector in the 1960s may have been able to do some good. But much has changed. As Elizabeth Warren

¹⁸ For example, a regulator can prosecute surgeons who order treatments that are obviously unnecessary, from the perspective of the regulator. *See, e.g.*, *United States v. Campbell*, 845 F.2d 1374, 1375 (6th Cir. 1988) (prosecuting a doctor for defrauding Medicare by ordering superfluous treatment).

¹⁹ *See* Hadfield et al., *supra* note 15, at 134 (distinguishing between informational and substantive regulation).

²⁰ *Schloendorff v. Soc’y of N.Y. Hosp.*, 105 N.E. 92, 93 (N.Y. 1914).

²¹ *See* Richard A. Epstein, *Regulatory Paternalism in the Market for Drugs: Lessons from Vioxx and Celebrex*, 5 *YALE J. HEALTH POL’Y L. & ETHICS* 741, 746–47 (2005) (“The regulator who works upstream of the physician and patient lacks any knowledge of individuated circumstances that should rationally influence the decision of which drug, if any, to take, and in what dosage. So long as physicians and patients have some skill in locating the patient’s position in the distribution, there is no reason to rely on the upstream averages that the FDA uses. Patients and physicians should be allowed to incorporate downstream knowledge into their decisions.”).

²² Colin Camerer et al., *Regulation for Conservatives: Behavioral Economics and the Case for “Asymmetric Paternalism,”* 151 *U. PA. L. REV.* 1211, 1212 (2003) (“[T]o the extent that paternalism prevents people from behaving in their own best interests, paternalism may prove costly.”).

writes, “[I]nnovation in financial products has produced incomprehensible terms and sharp practices that have left families at the mercy of those who write the contracts.”²³ When regulators do impose substantive controls, the financial industry simply innovates again to create new mechanisms to exploit their financial interests, in a pattern that scholars call a “regulatory dialectic.”²⁴ Whether industry moves through loopholes left by captured regulators, or by redefining the financial products into new fungible forms, the problems seem to just return.²⁵

Thus, real reform of health care and lending practices, to protect consumers and stabilize the economy over the long term, may require reform of the epistemic and economic situations in which patients and borrowers make their decisions. It may be more fruitful to focus reform efforts on those micro-level individual decisions themselves, if those are, after all, a root cause of the macro-level problems. There are several avenues for such regulation of the advisory relationship.

As an initial solution to this problem of biased advice, policymakers have mandated that advisors disclose their conflicting interests to the laypersons who rely upon them.²⁶ Reflecting this first policy solution is Federal Rule of Civil Procedure 26(a), which requires expert witnesses to disclose how much litigants pay for their services,²⁷ presumably so the laypersons on the jury can discount the testimony accordingly. Similarly, SEC Rule 10(b) requires that a broker who is acting as a principal in a transaction must disclose that fact to the customer.²⁸ Laws increasingly require that physicians disclose their ties to the pharmaceutical industry, at least indirectly through websites that are

²³ Warren, *supra* note 10, at 9.

²⁴ See e.g., Edward J. Kane, *Impact of Regulation on Economic Behavior: Accelerating Inflation, Technological Innovation, and the Decreasing Effectiveness of Banking Regulation*, 36 J. FIN. 355, 355 (1981); Merton H. Miller, *Financial Innovation: The Last Twenty Years and the Next*, 21 J. FIN. & QUANTITATIVE ANALYSIS 459, 461 (1986).

²⁵ See Kane, *supra* note 24, at 355; see also Nathalie Martin, *1,000% Interest—Good While Supplies Last: A Study of Payday Loan Practices and Solutions*, 52 ARIZ. L. REV. 563, 590 (2010) (identifying ways in which payday lenders tweaked or repacked their financial products to avoid consumer protection regulations).

²⁶ See Margaret Z. Johns, *Informed Consent: Requiring Doctors to Disclose Off-Label Prescriptions and Conflicts of Interest*, 58 HASTINGS L.J. 967, 1011–12, 1020–22 (2007) (detailing requirements that doctors disclose conflicting interests to patients).

²⁷ FED. R. CIV. P. 26(a).

²⁸ 17 C.F.R. § 240.10b–10 (2009).

theoretically available to patients.²⁹ Realtors who undertake to represent both the buyer and seller in a transaction are required to notify the clients that “[r]epresenting more than one party to a transaction presents a conflict of interest since both clients may rely upon [the realtor’s] advice and the client’s respective interests may be adverse to each other.”³⁰ Such disclosure mechanisms can serve two purposes: protecting laypersons’ autonomy to make informed choices and improving the quality of the choices they make.³¹ However, recent economic modeling and empirical research suggest that disclosure mandates may be counterproductive to the layperson’s own welfare if they worsen the quality of advice given or undermine trust, and yet fail to improve layperson performance.³² Still, many laypersons say that they want disclosures,³³ and policymakers continue to institute new and broader disclosure mandates.³⁴

Another policy response is to proscribe the conflict by banning those who advise laypersons from also having conflicting interests.³⁵ For example, federal law prohibits doctors from receiving kickbacks for referring patients.³⁶ Likewise, the FDA “permits financially *disinterested* physicians to promote off-label indications . . . but forbids other physicians” who have ties to the

²⁹ Arlene Weintraub, *New Health Law Will Require Industry to Disclose Payments to Physicians*, KAISER HEALTH NEWS (Apr. 26, 2010), <http://www.kaiserhealthnews.org/stories/2010/april/26/physician-payment-disclosures.aspx> (describing various laws that require physician–patient disclosure).

³⁰ ILL. ASS’N OF REALTORS, FORM 335: DISCLOSURE AND CONSENT TO DUAL AGENCY (2000), available at <http://www.ppreservices.com/forms/dualagencyconsent.pdf>.

³¹ Dennis F. Thompson, *Understanding Financial Conflicts of Interest*, 329 NEW ENG. J. MED. 573, 575 (1993) (“An advantage of disclosure is that it gives those who would be affected, or who are otherwise in a good position to assess the risks, information they need to make their own decisions.”); see also Johns, *supra* note 26, at 1015; Marc A. Rodwin, *Physicians’ Conflicts of Interest*, 321 NEW ENG. J. MED. 1405, 1406 (1989).

³² Daylian M. Cain, George Loewenstein & Don A. Moore, *The Dirt on Coming Clean: Perverse Effects of Disclosing Conflicts of Interest*, 34 J. LEGAL STUD. 1, 18 (2005); Ming Li & Kristóf Madarász, *When Mandatory Disclosure Hurts: Expert Advice and Conflicting Interests*, 139 J. ECON. THEORY 47, 48–50, 60, 62–63 (2008).

³³ Christine Grady et al., *The Limits of Disclosure: What Research Subjects Want to Know About Investigator Financial Interests*, 34 J.L. MED. & ETHICS 592, 597–98 (2006).

³⁴ Troyen A. Brennan & Michelle M. Mello, *Sunshine Laws and the Pharmaceutical Industry*, 297 JAMA 1255, 1256 (2007).

³⁵ See Troyen A. Brennan et al., *Health Industry Practices That Create Conflicts of Interest*, 295 JAMA 429, 431 (2006) (“[M]any current practices should be prohibited and others should be more strictly regulated to eliminate potential sources of unwarranted influence.”).

³⁶ 42 U.S.C. § 1320a–7b(b) (2006); see also, e.g., *United States v. Goss*, 96 F. App’x 365 (6th Cir. 2004) (applying the anti-kickback statute in the context of diagnostic referrals).

pharmaceutical industry from undertaking those same promotions.³⁷ In the field of human subjects research, where clinicians induce their patients to join clinical trials, there are widespread calls for additional limits to create “financial neutrality between treatment and research, thus ensuring that a physician’s decision to conduct clinical research, as well as his or her decision to recommend that a particular individual participate in a clinical trial, is grounded in reasons unrelated to investigator compensation.”³⁸ If these policies succeed, they convert conflicted advisors into non-conflicted advisors. They do so by forcing the advisor to choose between her advisory business and her alternative source of business.

An alternative policy option concedes that primary advisors may be biased but mandates that particularly vulnerable laypersons be given independent unbiased advice, before acting on the advice provided by conflicted experts. Some states require that senior citizens get a second opinion from an independent advisor before agreeing to a reverse mortgage on their homes.³⁹ Medicaid, Medicare, and private health insurers have required that patients get second opinions before acting on advice from physicians with conflicting interests.⁴⁰ Likewise, under Oregon’s Death with Dignity Act, a treating physician may recommend assisted suicide, but a patient seeking to end her life must get confirmation from a consulting physician, who may approach the case more objectively.⁴¹ The federal requirement that clinical research studies using human subjects must first be approved by an institutional review board (IRB) may also reflect this insight, because one primary function of the IRB is to independently assess the risks to layperson participants and provide some advice about those risks in an “informed consent” form.⁴² There are also

³⁷ Gregory Conko, *Truth or Consequences: The Perils and Protection of Off-Label Drug and Medical Device Promotion*, 21 HEALTH MATRIX (forthcoming 2011) (manuscript at 15), available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1677609.

³⁸ KATHLEEN M. BOOZANG ET AL., SETON HALL UNIV. SCH. OF LAW, CTR. FOR HEALTH & PHARM. LAW & POLICY, CONFLICTS OF INTEREST IN CLINICAL TRIAL RECRUITMENT & ENROLLMENT: A CALL FOR INCREASED OVERSIGHT 1 (2009), available at http://law.shu.edu/ProgramsCenters/HealthTechIP/upload/health_center_whitepaper_nov2009.pdf.

³⁹ E.g., MASS. ANN. LAWS ch. 167E, § 7 (LexisNexis 2009).

⁴⁰ Susan P. Shapiro, *Bushwhacking the Ethical High Road: Conflict of Interest in the Practice of Law and Real Life*, 28 LAW & SOC. INQUIRY 87, 238 (2003); see, e.g., *Damare v. Occidental Petroleum Corp. Med. Care Plan*, No. 92-1779, 1993 WL 92503, at *3 (E.D. La. Mar. 24, 1993) (quoting the second-opinion policy of one health insurer).

⁴¹ OR. REV. STAT. § 127.820 (2010).

⁴² See 21 C.F.R. § 56.109 (2010).

various ombudsperson programs, in which a purportedly independent advisor is assigned to protect the interests of a vulnerable class of persons.⁴³

Some policies nudge laypersons toward independent advice, without actually mandating it. For example, lawyers are prohibited from entering into business transactions or settling malpractice claims with their own clients, unless the client is first “advised in writing of the desirability of seeking and is given a reasonable opportunity to seek the advice of independent legal counsel on the transaction.”⁴⁴ Likewise, realtors who propose to serve in dual agency relationships must advise their clients “to seek independent advice from [their] advisors or attorneys before signing any documents in this transaction.”⁴⁵ This sort of policy is something more than a disclosure of a conflict, but less than a mandate for a second opinion.

A related policy response is for the regulator itself to provide independent advice, or at least user-friendly information, to laypersons. In the litigation setting, courts have long had the power to bring their own expert witnesses, as an antidote to the biases of hired-gun expert witnesses.⁴⁶ With few exceptions, the courts have generally declined to do so, however.⁴⁷ In the market, government-mandated vehicle rollover ratings, gas mileage ratings, appliance efficiency standards, and annual percentage rates can be useful alternatives to the cheap talk of a salesman.⁴⁸ These interventions can be viewed as providing alternative sources of unbiased advice, or they can be understood as more fundamental solutions that reduce the level of epistemic asymmetry between advisor and client, by raising the abilities of the client.

Another policy solution is to do nothing, to assume that the market will itself resolve this problem. If laypersons need unbiased advice to make

⁴³ See, e.g., 11 U.S.C. § 333 (2006) (Bankruptcy Code provision providing for appointment of “patient care ombudsman” when health care provider declares bankruptcy); DEP’T OF HEALTH & HUMAN SERVS., EFFECTIVE OMBUDSMAN PROGRAMS (1991), available at <http://oig.hhs.gov/oei/reports/oei-02-90-02122.pdf> (surveying six such programs in the nursing home context); Maxwell J. Mehlman, *Medical Advocates: A Call for a New Profession*, 1 WIDENER L. SYMP. J. 299 (1996) (describing such programs in nursing homes and managed care programs).

⁴⁴ MODEL RULES OF PROF’L CONDUCT R. 1.8(a)(2), (h)(2) (2010).

⁴⁵ ILL. ASS’N OF REALTORS, *supra* note 30, at 1.

⁴⁶ FED. R. EVID. 706.

⁴⁷ See Christopher Tarver Robertson, *Blind Expertise*, 85 N.Y.U. L. REV. 174, 199–201 (2010).

⁴⁸ See generally ARCHON FUNG ET AL., FULL DISCLOSURE: THE PERILS AND PROMISE OF TRANSPARENCY (2007) (reviewing the history of informational disclosure mandates). On the other hand, in the financial industry, the pages and pages of legalese disclosures seem to simply present an opportunity to hide the most unscrupulous needles in a haystack of verbiage. See Warren, *supra* note 10, at 11–12 (describing the increasing length and complexity of credit card contracts).

decisions for their own welfare, then there should be a market of such advisors; laypersons could simply buy the advice that they need, paying a premium for unbiased over biased advisors, if necessary. For example, given the inaction of the courts to address the hired-gun problem in litigation, I have developed the concept of "blind experts," which would be brought by litigants themselves acting in their own self-interests.⁴⁹ Or in the health care setting, if one is concerned about the conflicts of interest inherent in a fee-for-surgery practice, one can instead join a managed care organization, though these surgeons may have the opposite biases.⁵⁰ In the financial markets, there are brokerages who are compensated on a fee-per-trade basis (which thus creates an incentive to churn the accounts), and there are others compensated on the basis of the amount of assets under management (which thus creates an incentive to perform, or to invest money in advertising for more clients at least).⁵¹ There are other tradeoffs to be made; it may not be possible to perfectly align incentives, and laypersons may fail to appreciate and appropriately value non-conflicted advice over conflicted advice.⁵² Whether laypersons actually do so is an empirical question explored below.

So a range of potential policy responses exists. Unfortunately, the comparative effectiveness of these multifarious policy alternatives remains poorly understood. Through a series of behavioral experiments in a laboratory setting, the present study tests these policies against each other and advances the hypothesis that the production and provision of unbiased sources of advice is the most promising policy solution to this problem of biased advice in contexts of epistemic asymmetry.

⁴⁹ Robertson, *supra* note 47, at 179–80.

⁵⁰ See Howard Brody, *The Physician–Patient Relationship*, in MEDICAL ETHICS 75, 93 (Robert M. Veatch ed., 2d ed. 1997) (describing conflicts between patient welfare and obligations to health care plans in managed care situations).

⁵¹ Craig J. McCann, *Churning*, 9 J. LEGAL ECON. 49, 49 (1999).

⁵² See generally Saul Levmore, *Commissions and Conflicts in Agency Arrangements: Lawyers, Real Estate Brokers, Underwriters, and Other Agents' Rewards*, 36 J.L. & ECON. 503 (1993) (explaining why solutions to these sorts of agency problems are not found in practice as frequently as one might expect based on economic theory).

I. HOW A MANDATORY DISCLOSURE POLICY CAN HURT LAYPERSONS BY
DEGRADING THE ADVICE GIVEN

A. *The Cain, Loewenstein, and Moore Study (CLM)*

Only recently have scholars begun to test empirically how mandated disclosures about experts' conflicting interests actually impact layperson decision making. One might worry that such disclosure policies are useless, as several studies have suggested.⁵³ However, in 2005, Daylian Cain, George Loewenstein, and Don Moore published a study (CLM) concluding that disclosure mandates can actually be deleterious.⁵⁴ A disclosure mandate may actually hurt the very laypersons it is designed to protect.⁵⁵

The CLM study merits extended discussion here not only for its intrinsic interest, but also because its methods are the basis for the present study. CLM put students at Carnegie Mellon University in either of two roles, "estimators" and "advisors," with the task of ascertaining the values of assorted coins in each of six jars.⁵⁶ This estimation task served as a proxy for real-world tasks that laypersons face, such as deciding how much a house is worth, how much a company stock is worth, and whether a surgical procedure is worthwhile given

⁵³ For example, in one survey-based study, Lindsay Hampson and colleagues found that "[m]ost patients in cancer-research trials were not worried about financial ties between researchers or medical centers and drug companies and would still have enrolled in the trial if they had known about such financial ties." Lindsay A. Hampson et al., *Patients' Views on Financial Conflicts of Interest in Cancer Research Trials*, 355 NEW ENG. J. MED. 2330, 2330 (2006). An experimental study by Kevin Weinfurt and colleagues randomized human subjects considering whether to participate in a hypothetical clinical trial into three conditions: one where there was no disclosed conflict, one where the researchers disclosed that they had an equity stake in an interested business, and one where the researchers disclosed receiving a per-participating-patient payment from an interested business. Kevin P. Weinfurt et al., *Effects of Disclosing Financial Interests on Participation in Medical Research: A Randomized Vignette Trial*, 156 AM. HEART J. 689 (2008). Subjects in the equity group expressed significantly less willingness to participate than in the other two conditions, though the causal mechanism for this preference between the two forms of conflict was unclear. *Id.* at 691. Since there was no way to specify the optimal participation rate in each condition, the Weinfurt study provides no way to assess whether, on net, the disclosure mandate helped or hurt the participants.

The disclosure problem also arises at a higher level, where physicians are the relative laypersons relying on the expertise of scientists advising them through biomedical journal articles. Gabriel Silverman and colleagues tested physicians reviewing biomedical journal abstracts that reported the efficacy of a new drug, with and without disclosed conflicts of interest. Gabriel K. Silverman et al., *Failure to Discount for Conflict of Interest When Evaluating Medical Literature: A Randomised Trial of Physicians*, 36 J. MED. ETHICS 265 (2010). The study found that the disclosures had no significant impact on the physicians' reliance on the study, as measured by the physicians' likelihood of prescribing the drug. *Id.* at 265.

⁵⁴ Cain et al., *supra* note 32.

⁵⁵ *Id.* at 22.

⁵⁶ *Id.* at 9.

its apparent benefits and costs.⁵⁷ Although contrived and stylized, the coins task allowed the researchers to specify a concrete measure of accuracy, and thus provided a mechanism for judging layperson performance that may be analogous to real-world measures of utility (such as health or wealth), where the layperson's practical decision turns out to be objectively good or bad for him.

To create epistemic asymmetry, CLM gave the estimators only glimpses of the jars of coins at a distance, but the advisors were given some expertise in the task, as they had more time to hold and examine the jars and were told a range of potential values.⁵⁸ CLM also created conflicting interests. The CLM estimators were always compensated on the basis of the accuracy of their estimates, while the advisors' compensation varied across the three conditions of the study.⁵⁹ In the first condition (labeled "accurate"), the advisors were compensated based on the accuracy of the estimators, thus aligning their interests, and the estimator was advised of this fact.⁶⁰ In the second ("high/disclosed") and third ("high/undisclosed") conditions, the advisors were told that they would be compensated based on how high the estimator's guess was. This fact was disclosed to the estimators in the second, but not the third, condition, and the advisors knew whether their conflict would be disclosed.⁶¹ Thus, CLM was able to test the comparative effectiveness of the disclosure mandate in the high/disclosed condition versus the high/undisclosed condition, to determine which one best approximated the performance of the accurate condition.

CLM found that estimators performed best in the accurate condition and somewhat worse when receiving biased advice in the high/undisclosed condition, as would be expected.⁶² More surprisingly, across the two conditions where a conflict of interest existed, the estimators did *worse* in the mandatory disclosure condition (high/disclosed).⁶³ This occurred for two reasons. First, the advisors gave significantly worse advice in the disclosed

⁵⁷ *Id.* at 20.

⁵⁸ *Id.* at 9–10.

⁵⁹ *Id.* at 10.

⁶⁰ *Id.* After receiving the substantive advice, the estimators were told: "Note: The advisor is paid based on how accurate the estimator is in estimating the worth of the jar of coins." *Id.*

⁶¹ *Id.* The conflict was disclosed as follows: "Note: The advisor is paid based on how high the estimator is in estimating the worth of the jar of coins." *Id.* No such disclosure was given in the third condition, even though there was a conflicting interest. *Id.*

⁶² *Id.* at 17.

⁶³ *Id.*

condition than in the undisclosed condition.⁶⁴ The advisors apparently felt that the disclosure gave them a “moral license” to be even more biased, since the layperson was on notice that the advice may be biased and they could take it or leave it.⁶⁵ *Caveat emptor*. Second, the estimators failed to effectively use the disclosure to adjust for the inaccuracy of the given advice, presumably because they had little independent way to assess whether and to what extent the advisors were actually biased and because they had no other source of advice to rely upon instead.⁶⁶ When you are told that your only advisor is conflicted, it is not precisely clear what you should do with such information.

B. The Present Experiment’s Replication and Extension of CLM

Like the CLM study, the present study involved layperson estimators relying on advisors for a coins-in-jars estimation task with incentives for accuracy, but this Article’s study was conducted online. The specific methods for recruiting human subjects and running the experiment are described in the notes and the Methodological Appendix.⁶⁷ The study focused only upon the behavior of the estimators, here called laypersons, across twelve experimental conditions. Unbeknownst to the participants, the expert advice was simulated based on the results of the CLM study, using the means for reported advice given in the accurate, high/disclosed, and high/undisclosed conditions.⁶⁸

Table 1 in the Appendix presents results on the comparable conditions in the CLM study and the present study. To measure the effectiveness of disclosures, the CLM study and the present study used “virtual errors,” which are defined as the absolute value of the difference between the layperson’s estimate and the expert’s personal assessment in the accurate condition,

⁶⁴ *Id.* at 13.

⁶⁵ *Id.* at 7.

⁶⁶ *Id.* at 17.

⁶⁷ This study used the same values of coins in six jars as the CLM study, though subjects were shown small, low-resolution photographs of such jars rather than actual jars. The jars photographed had the same total value as each of the CLM study’s jars but likely consisted of different combinations of quarters, dimes, nickels, and pennies. The photographs were 359 by 336 pixels in size. See the Methodological Appendix *infra* for more information about the jars’ values and photographs thereof.

⁶⁸ As in the CLM study, participants were told that they would receive advice from “advisors who have actually held those jars, who had several minutes to examine them, and who have been told the range of potential values.” *Cf.* Cain et al., *supra* note 32, at 9 (discussing experimental methods). The participants were given the advice and other prompts depending on the experimental conditions, and were then asked to render estimates of the value of the coins. After each estimate, the laypersons also disclosed their confidence in the accuracy of their estimates. Once answers were submitted for one jar, participants then repeated the task for another jar and were not able to go back and change their answers.

averaged across the six jars.⁶⁹ This provides, as a benchmark, a measure of what an independent, well-informed observer thinks. If the laypersons performed as well as an expert, then one might assume that the advisory relationship was working perfectly.⁷⁰

It is worthwhile to attempt to replicate the CLM findings.⁷¹ Doing so confirms the essential relationships shown by the CLM study. First, as one would expect, laypersons performed worse when receiving biased advice. In the high/undisclosed condition (called 1BN here, for one expert who is biased but with no disclosure), their errors were larger than when relying upon unbiased advisors in the accurate condition (called 1UA here)—a difference of \$1.29, or 36%.⁷² More interestingly, just as in the CLM study, laypersons with biased advisors but no disclosures (1BN) did much better than those who had biased advisors who gave mandated disclosures of the conflict (1BC)—a difference of \$1.64, or 34%.⁷³ All of the point estimates in the present study are statistically indistinguishable from those in the CLM study.⁷⁴

⁶⁹ See Cain et al., *supra* note 32, at 13 n.7 (defining virtual error); *id.* 16 tbl.6 (disclosing estimators' personal estimates).

⁷⁰ Still, a more obvious dependent variable would be to measure the absolute value of the difference between the layperson's estimate and the true value, and these results for "actual error" are reported in the Appendix. Following the CLM study, virtual error is instead used in this Article to account for the fact that both laypersons and experts systematically underestimated the value of the coins in the jars, and to avoid the peculiar finding that the advisor's upward bias due to a conflicting interest does not harm layperson accuracy, but instead helps correct for the natural bias. When absolute error is analyzed rather than virtual error, both the CLM study and the present studies found no statistical difference in *layperson* performance between those with accurate advisors and those with undisclosed conflicted advisors, even though the *advisors* gave significantly more biased advice in the latter condition. The difference in means is only \$0.07 in the present study ($p = .94$). Although this is certainly a possible circumstance in real-world situations of epistemic asymmetry with conflicting interests, this would be a special case, and the study has greater external validity once that anomaly is resolved by reference to virtual error instead. Thus, henceforth this Article simply uses *layperson inaccuracy* as the primary dependent variable, but refers to virtual error in doing so.

⁷¹ Ramal Moonesinghe et al., *Most Published Research Findings Are False—But a Little Replication Goes a Long Way*, 4 PLoS MED. e28 0218, 0218 (2007), [http://www.plosmedicine.org/article/ fetchObjectAttachment.action?uri=info%3Adoi%2F10.1371%2Fjournal.pmed.0040028&representation=PDF](http://www.plosmedicine.org/article/fetchObjectAttachment.action?uri=info%3Adoi%2F10.1371%2Fjournal.pmed.0040028&representation=PDF) ("As part of the scientific enterprise, we know that replication—the performance of another study statistically confirming the same hypothesis—is the cornerstone of science and replication of findings is very important before any causal inference can be drawn.")

⁷² $M_{1BN} = 4.85$ ($SE = 0.40$), $M_{1UA} = 3.56$ ($SE = 0.42$), $t(80) = 2.20$, $p = .03$, $r = .24$; see Cain et al., *supra* note 32, at 16 tbl.6 (reporting this data from the CLM study); *infra* Table 1 (providing statistical comparisons).

⁷³ $M_{1BN} = 4.85$ ($SE = 0.40$), $M_{1BC} = 6.49$ ($SE = 0.30$), $t(157) = -2.93$, $p < 0.01$, $r = .23$; see Cain et al., *supra* note 32, at 16 tbl.6 (reporting this data from the CLM study); *infra* Table 1 (providing statistical comparisons).

⁷⁴ See *infra* Table 1 (reporting statistical comparisons).

As in CLM, a mandatory disclosure policy does not seem to help laypersons adjust their reliance on the advice received. Instead, it may only cause the expert advisors to become more biased.⁷⁵ Policymakers should thus be wary about the value of the disclosure mandates as a solution to conflicting interests.

Further study is necessary to understand whether and how to improve disclosure policies, and to explore alternative policy mechanisms to help laypersons in these situations of epistemic asymmetry and conflicting interests. The researcher fielded nine other experimental conditions for this purpose.⁷⁶ These conditions are discussed in the Parts that follow.

II. WHEN A DISCLOSURE, OR EVEN A BAN, MIGHT WORK, DEPENDING ON RELATIVE EXPERTISE AND DEGREE OF BIAS

A. *Measuring Epistemic Asymmetry and Bias*

A layperson–advisor relationship involves two distinct factors that impact layperson performance in context-dependent ways. First is the degree to which the advisor has expertise compared to the layperson, and second is the degree to which the advisor is subject to biases caused by conflicting interests. Each of these dimensions must be accounted for in policy making and experimental design.

The first factor is the difference between the estimation skills of the estimator (given his situation) and the advisor (given her situation); the advisor’s comparative expertise is the very reason why the layperson may be tempted to place his reliance on the advisor. Alternatively, this factor could be called “epistemic asymmetry.”⁷⁷ In the law of evidence the notion of being an “expert” is defined by a witness having “knowledge, skill, experience, training, or education” that the layperson jurors lack, and which would “assist” the jury in deciding the case.⁷⁸ In some situations, there will be a great disparity

⁷⁵ To emphasize, this study does not retest the performance of the advisors (instead only assuming that they will perform as they did in the CLM study), but does replicate the findings showing how laypersons react to disclosures of conflicted interests.

⁷⁶ See *infra* Table 3.

⁷⁷ The term *information asymmetry* is widely used in economic bargaining theory. See, e.g., Hadfield et al., *supra* note 15. *Epistemic asymmetry* is somewhat broader, since it also includes the skill, experience, training, or education that allows a party to make practical sense of the information that may be available to that party.

⁷⁸ FED. R. EVID. 702.

between the skills of the layperson and the advisor, who is truly an expert. In other situations, the advisor will have no real epistemic advantage. For example, expert testimony is not necessary to prove that a surgeon should remove his instruments and surgical sponges before sewing up a patient.⁷⁹ As the Federal Rule of Evidence 702 Advisory Committee noted:

There is no more certain test for determining when experts may be used than the common sense inquiry whether the untrained layman would be qualified to determine intelligently and to the best possible degree the particular issue without enlightenment from those having a specialized understanding of the subject involved in the dispute.⁸⁰

Epistemic asymmetry is thus a relative measurement.

The same is true for conflicting interests. There are two potential concerns with conflicting interests—they can create biases in the advice given, and they can decrease the layperson's trust in his advisor, if the bias is disclosed or observed. For now, let us focus on the first problem.⁸¹ There will be cases in which the conflicted expert has such extreme biases that his opinion will be almost worthless, even if he is highly skilled.⁸² In other cases, the conflicted expert will have no discernable biases and thus be quite likely to provide his best estimate.

Thus, expertise and bias are two different dimensions of accuracy. Measurement of epistemic asymmetry was not possible given the design of the CLM experiment,⁸³ but the present study allows such measurement and thus allows more calibrated policy recommendations. In the present study, condition NoAdvisors asked the layperson to perform the estimation task without any expert advice at all. Laypersons in the NoAdvisors condition erred by \$11.65 on average.⁸⁴ CLM reports that, in the accurate condition, the

⁷⁹ See, e.g., *Burke v. Wash. Hosp. Ctr.*, 475 F.2d 364, 366 (D.C. Cir. 1973) (explaining that when a surgeon leaves his tools in a patient, it “appears to be that rare sort of case in which the type of harm itself raises so strong an inference of negligence, and the physician’s duty to prevent the harm is so clear, that expert testimony is not required to establish the prevailing standard of care”).

⁸⁰ FED. R. EVID. 702 advisory committee’s note (quoting Mason Ladd, *Expert Testimony*, 5 VAND. L. REV. 414, 418 (1952)) (internal quotation marks omitted).

⁸¹ The latter point is explored in Part IV *infra*. It is also worth noting that conflicting interests are not the only source of biases. Other biases are beyond the scope of this paper.

⁸² See, e.g., *In re Silica Prods. Liab. Litig.*, 398 F. Supp. 2d 563, 627–28, 640 (S.D. Tex. 2005) (excluding expert testimony in part because compensation bias was prominent).

⁸³ Cain et al., *supra* note 32, at 16 tbl.6 (showing that lack of NoAdvisors condition in the CLM study).

⁸⁴ See *infra* Table 4 (reporting actual errors rather than virtual errors).

advisors personally estimated that the jars held \$15.62 on average,⁸⁵ but the jars actually held \$18.16 on average,⁸⁶ which means that the experts themselves erred by \$2.54 on average (the difference). When one compares this \$2.54 actual expert error to the \$11.65 actual error of laypersons, we can compare the expertise of these two actors and see the epistemic asymmetry. Dividing these two average actual errors, one can conclude that in this experimental setting there is an epistemic asymmetry ratio of 459% between experts and laypersons. The errors of laypersons were more than four times the size of those of the unbiased experts.

One can likewise calculate a “bias ratio” to capture the inaccuracy of the advice offered when the advisor has interests aligned with the estimator, compared to when those interests are conflicted. First, to compute the bias when interests are aligned, subtract the advisors’ personal estimates (\$15.62) in the accurate condition from the average proffered advice (\$16.48) in that same condition; this yields \$0.86.⁸⁷ Second, to compute the bias when interests are conflicted, subtract the advisors’ personal estimates in the accurate condition (again \$15.62) from the average proffered advice in the high/undisclosed condition (\$20.16); this yields \$4.54.⁸⁸ As one can see, the discrepancy between the proffered advice and what advisors actually believe (that is, their personal estimates) increases from \$0.86 to \$4.54, when interests shift from aligned to conflicted. Dividing \$4.54 by \$0.86, we compute a bias ratio of 528%, which is the degree to which the inaccuracy of advice increases when interests are conflicted rather than aligned. In other words, in this study, advisors with conflicting interests give advice that is more than five times as inaccurate as advisors with aligned interests.

Thus, the CLM study and the present study explore a situation of large epistemic asymmetry of 459% and large bias of 528%. CLM found that a disclosure mandate did not help in this setting,⁸⁹ and this further analysis suggests that the reason may be that a layperson who rejected the biased advice would be left with his own poor estimates. Thus, regardless of whether the laypersons followed the bad advice or trusted their own bad estimates, the

⁸⁵ Cain et al., *supra* note 32, at 15 tbl.5.

⁸⁶ *Id.* at 14 tbl.4 (averaging across row 1).

⁸⁷ *See id.* at 15 tbl.5. It is not clear why there was any discrepancy between personal estimates and advice given in the condition where interests were aligned. It is possible that advisors were trying to offset systematic errors that they presumed that their estimators might make.

⁸⁸ *See id.*

⁸⁹ *Id.* at 6–7.

result was unlikely to be very good. This is the classic “out of the frying pan into the fire” sort of problem.

So, these ratios show that the estimates of the advisor and the layperson were both bad, just for different reasons. In this context, the disclosure mandate simply made the problem worse, since it worsened the advice given even further. The disclosure mandate essentially imposed a transaction cost on those laypersons who used the disclosure to switch their reliance from the advisor to themselves, with no real benefit.

B. Extrapolating to Real World Conditions to Test Policy Solutions

Can we generalize from the CLM study? It is important to emphasize that CLM constructed an artificial experiment in which the researchers created experts, who actually had privileged epistemic access to the truth (the value of coins in a jar). Yet, in the real world, not every advisor is an expert. Indeed, the CLM setting may be more the exception than the rule.⁹⁰ So, before extrapolating these findings, it would be useful to have a measure of the expertise ratio and the bias ratio in the specific setting where a disclosure mandate is proposed. Only if the ratios are comparable to those tested in the CLM study should we expect that the laboratory findings will have predictive value. What about real-world situations where “the experts” do not actually have much expertise? Or where the conflicting interests do not actually create biases? This section explores those variations.

Take the doctor–patient relationship. Plausibly, one might suppose that the epistemic asymmetry in the typical doctor–patient relationship is quite high (perhaps more than 459%), given the hard science underlying much of medicine, the extensive formal training physicians receive, and their individual and collective experience.⁹¹ As for the bias ratio, one might hope that the

⁹⁰ See generally DAVID H. FREEDMAN, *WRONG 7* (2010) (“The fact is, expert wisdom usually turns out to be at best highly contested and ephemeral, and at worst flat-out wrong.”).

⁹¹ Carl E. Schneider & Mark A. Hall, *The Patient Life: Can Consumers Direct Health Care?*, 35 *AM. J. L. & MED.* 7, 31–34 (2009). Still, there are contexts where physicians have very little hard evidence to go on and may be proceeding on little more than trial and error. See, e.g., Kevin A. Kerber & A. Mark Fendrick, *The Evidence Base for the Evaluation and Management of Dizziness*, 16 *J. EVALUATION CLINICAL PRAC.* 186, 189 (2010) (“Physicians rely on the medical literature to inform decisions, but our study suggests that the evidence base for dizziness evaluation and management is weak.”); Christian Davenport, *Doctors Who Prescribe Off-Abused Drugs Face Scrutiny*, *WASH. POST*, Jan. 1, 2011, at A01 (“Doctors ‘don’t get very much, if any, training in dependence, in addiction, in pain management’ . . .” (quoting R. Gil Kerlikowske, Director, White House Office of National Drug Control Policy)). And, evidence suggests that patients are increasingly turning to their own epistemic resources (such as WebMD or nontraditional healers), which may make the epistemic

professionalism of doctors will minimize the size of the financial biases in their advice, making it much smaller than CLM's observed 528%. Still, there is evidence that doctors (like all humans) respond to incentives, and incentives are often misaligned.⁹² Moreover, even unbiased doctors may render biased advice if it is based on scientific findings that are themselves biased by the pharmaceutical industry.⁹³

Therefore, in the particular setting of medical practice, CLM's findings may have relevance for policymakers—if the epistemic asymmetry and bias ratios are comparable. In other settings, where the epistemic asymmetry is smaller (because the advisors have little relative expertise), and the bias ratio is the same or larger (for example, if the advisor has few legal or social constraints on exploitive behavior), a disclosure mandate may be salutary. In that context, a disclosure mandate may cause laypersons to reject the biased advice and follow their own judgments instead.

Consider other contexts where it may be tempting to apply CLM's findings. One might suppose that the epistemic asymmetry in the retail stockbroker–investor relationship is relatively low for the task of selecting a stock or mutual fund, given the empirical research showing that performance of any particular investment is rarely better than random and almost impossible to predict.⁹⁴ Indeed, the efficient market hypothesis suggests that a random

asymmetry better or even worse, depending on the quality of that information and the layperson's ability to use it. See, e.g., Lisa Grossman, *The Net Doctor Will See You Now*, NEW SCIENTIST, July 25, 2009, at 20 (describing the increasing use of online medical resources in advance or in lieu of seeing a doctor). In principle, this overall asymmetry could be measured, for instance, by asking doctors and laypersons to each answer some context-relevant questions for which the answer is objective and scalar. For example, what is the one-year survival rate for patients with a given condition who go untreated? What is the one-year survival rate with the preferred treatment?

⁹² See sources cited *supra* notes 1–5.

⁹³ See generally *In re Zyprexa Prods. Liab. Litig.*, 253 F.R.D. 69, 106 (E.D.N.Y. 2008) (“The pervasive commercial bias found in today’s research laboratories means studies are often lacking in essential objectivity, with the potential for misinformation, skewed results, or cover-ups.”), *rev’d in part sub nom.* UFCW Local 1776 v. Eli Lilly & Co., 620 F.3d 121 (2d Cir. 2010); COMM. ON CONFLICT OF INTEREST IN MED. RESEARCH, EDUC. & PRACTICE, INST. OF MED., CONFLICT OF INTEREST IN MEDICAL RESEARCH, EDUCATION, AND PRACTICE 104 (Bernard Lo & Marilyn J. Field eds., 2009) (“Several systematic reviews and other studies provide substantial evidence that clinical trials with industry ties are more likely to have results that favor industry.”); Christopher T. Robertson, *The Triple Blind: How to Stop Industry Bias in Biomedical Science, Without Violating the First Amendment*, 37 AM. J.L. & MED. (forthcoming 2011) (reviewing the evidence of industry influence in biomedical science).

⁹⁴ See Laurent Barras et al., *False Discoveries in Mutual Fund Performance: Measuring Luck in Estimated Alphas*, 65 J. FIN. 179, 181–82 (2010) (examining performance of various mutual funds and finding that very few reliably beat the market); Andrew Metrick, *Performance Evaluation with Transactions Data: The Stock Selection of Investment Newsletters*, 54 J. FIN. 1743 (1999) (finding that the stock picks of

walk down Wall Street is likely to be just as effective, and surely less expensive, than hiring an advisor for advice.⁹⁵ In that sort of situation, the advisor and layperson will do roughly equally well. Nonetheless, a conflicted advisor may exert a strong bias toward frequent trades, churning the account to maximize transaction fees.⁹⁶

Likewise, real estate agents may have a relatively large bias toward advising their clients to buy (rather than rent) and pay more for a house, since the realtor is only paid upon a sale, and then as a portion of the sales price.⁹⁷ In these contexts, the bias of a conflicted advisor may be as large or larger than the 528% that can be derived from the CLM data.⁹⁸ Yet, the realtor might have very little real expertise for the task of predicting the appropriateness of a purchase for a particular family given its own needs and finances, nor will the realtor have any advantage in predicting *future* home prices.⁹⁹ One experimental study tasked both real estate agents and amateurs with appraising the market value of real houses, and found that both groups “were significantly biased by listing prices,” a factor which seems to beg the question about the true value of the house.¹⁰⁰ The authors noted that the agents seemed less aware of (or less candid about) the role of listing price in their estimates.¹⁰¹ Most importantly, the researchers found a “similarity of judgments” by both the experts and amateurs, and suggested that in such contexts where there appeared to be little epistemic advantage, “we might expect experts to talk a better game than amateurs, but to produce (on the average) similar results.”¹⁰²

investment newsletters fail to outperform the market). *But see* Kent L. Womack, *Do Brokerage Analysts' Recommendations Have Investment Value?*, 51 J. FIN. 137, 137 (1996) (analyzing data and concluding that stock “[a]nalytists appear to have market timing and stock picking abilities”).

⁹⁵ BURTON G. MALKIEL, *A RANDOM WALK DOWN WALL STREET: THE BEST AND LATEST INVESTMENT ADVICE MONEY CAN BUY* 24 (1996). *But see* Joshua D. Coval et al., *Can Individual Investors Beat the Market?* (Harvard Bus. Sch. Fin. Unit Research Paper Series, Working Paper No. 04-025, 2005), available at <http://ssrn.com/abstract=364000> (presenting evidence that some skillful individual investors do appear to reliably beat the market).

⁹⁶ *See* McCann, *supra* note 51, at 49; Roni Michaely & Kent L. Womack, *Conflict of Interest and the Credibility of Underwriter Analyst Recommendations*, 12 REV. FIN. STUD. 653 (1999) (presenting evidence that stock analysts are biased by their relationships to the companies they rate).

⁹⁷ Mark S. Nadel, *A Critical Assessment of the Traditional Residential Real Estate Broker Commission Rate Structure (Abridged)*, 5 CORNELL REAL EST. REV. 26, 33 (2006).

⁹⁸ *See supra* text accompanying notes 86–88.

⁹⁹ Nadel, *supra* note 97, at 39–40.

¹⁰⁰ Gregory B. Northcraft & Margaret A. Neale, *Experts, Amateurs, and Real Estate: An Anchoring-and-Adjustment Perspective on Property Pricing Decisions*, 39 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 84, 95 (1987).

¹⁰¹ *Id.*

¹⁰² *Id.* at 95–96.

In these contexts, where epistemic asymmetry is low and bias is high, CLM's findings may be inapposite. A disclosure mandate that informs the layperson of the conflicting interest and drives the layperson away from such advice may be salutary, especially if it is strengthened in the ways discussed below. The efficacy of a disclosure mandate is thus highly contingent on context, as measured by these two ratios. Indeed, as the epistemic asymmetry ratio approaches zero and the bias ratio grows, other policy interventions, such as an outright ban on those with conflicted interests providing advice, will become more salutary. If the conflicting interests cause large biases, but the advisor has very little epistemic advantages anyway, then the net advice is unlikely to be helpful.

The NoAdvisors condition shows that this is definitely not the case under the present experimental design borrowed from CLM. When the layperson has no advisors at all, the layperson errs by \$9.76 on average, which is much worse than the \$6.49 error in condition 1BC, where a layperson is given advice from one expert with a bias and a disclosure of conflicting interests.¹⁰³ Indeed, the errors in the NoAdvisor condition are significantly worse than any other condition. Under this experimental setting, biased advice is much better than nothing.

This huge difference in layperson performance suggests that in contexts of epistemic asymmetry that are similar to the one tested here, it may be much more important to ensure that laypersons have *some* advice than it is to worry about whether that advice is biased (or not) or whether that bias is disclosed (or not). For example, in some regions in the United States, there is a severe shortage of primary care physicians, and thus many laypersons are not getting the preventative care they need.¹⁰⁴ Such persons may not be receiving efficient and necessary treatments such as prescription statins, which are shown not only to help patients but also to reduce net health care costs.¹⁰⁵ One could imagine a policy in which the pharmaceutical companies that manufactured statins sent their own health care professionals into underserved areas with the

¹⁰³ $M_{\text{NoAdvisor}} = 9.76$ ($SE = 0.44$), $M_{\text{1BC}} = 6.49$ ($SE = 0.30$), $t(156) = 5.7$, $p < .001$, $r = .42$; see *infra* Table 3.

¹⁰⁴ Howard K. Rabinowitz et al., *Critical Factors for Designing Programs to Increase the Supply and Retention of Rural Primary Care Physicians*, 286 JAMA 1041, 1041 (2001) ("The shortage of primary care physicians in rural areas has been one of the most intractable US health policy problems of the past century.").

¹⁰⁵ Sheila Leatherman et al., *The Business Case for Quality: Case Studies and an Analysis*, 22 HEALTH AFF. 17, 20 (2003) ("Taking into account the clinical research literature on statins and statistical estimates of the longer-term costs of repeat heart attacks, the estimated ratio of cost to savings for effective treatment would be approximately 1:2.").

express goal of prescribing the drug, likely being biased in their decision making and thus overprescribing the drug compared to the optimal level. In such a context, if the cost of over-prescribing because of biased advice is less than the cost of underprescribing for lack of advice, policymakers might rationally prefer that laypersons receive such biased advice.

A ban on conflicted advice, on the other hand, can be dangerous in some contexts and helpful in others. Generally, where epistemic asymmetry is high, a ban on conflicted advice would be very bad policy, unless the policymaker can be confident that non-conflicted advisors would replace the conflicted advisors. Such replacement is not an obvious outcome of a ban on conflicted advice. To the extent that an advisor has a conflicting interest, the advisory services are being subsidized by some outside source.¹⁰⁶ Once that subsidy is removed by a ban policy, the layperson may no longer be able to afford the services of the advisor, who may instead find more lucrative work elsewhere. The conflict of interest may also be a function of the same relationship that creates the epistemic expertise. “For example, many both inside and outside the accounting industry have argued that an auditing firm is better equipped to handle a client’s complex accounting tasks when the auditor also has deep consulting ties to that client.”¹⁰⁷ Thus, policymakers must ask whether the asymmetry is greater than the bias ratio, and whether there is a viable alternative epistemic and economic relationship.

III. MAKING DISCLOSURES WORK BETTER THROUGH ANCHORING, INFORMATION TECHNOLOGY, AND PERSONALIZATION

Not all mandatory disclosures are created equal. This Part explores three potential ways to improve the efficacy of disclosures. First, policymakers might manipulate *when* disclosures are given, whether before or after the substantive advice. Second, policymakers might attempt to improve the *type* of disclosures given, to better enable laypersons to calibrate their advice. Third, policymakers can pay closer attention to *who* needs to receive disclosures, so as to maximize the benefits and minimize the harms of disclosure.

¹⁰⁶ William M. Sage, *Some Principles Require Principals: Why Banning “Conflicts of Interest” Won’t Solve Incentive Problems in Biomedical Research*, 85 TEX. L. REV. 1413, 1448–49 (2007).

¹⁰⁷ Moore et al., *supra* note 16, at 11.

A. When to Disclose

Prior behavioral research has shown that persons utilize an “anchor-and-adjust heuristic” to make decisions, one that is susceptible to undue influence from an initial prompt even after subsequent information is received.¹⁰⁸ If advice is provided first and a disclosure provided thereafter (as in CLM), the layperson may anchor on the bad advice before learning that it is unreliable. CLM speculated that such an anchoring problem may be a reason that disclosures fail.¹⁰⁹ Yet, this is a contingency that can be changed. I hypothesized that a disclosure mandate may work to improve layperson performance if the disclosure is given before rather than after the substantive advice. Condition 1BCF (one biased advisor with a conflict disclosed first) tests this hypothesis against condition 1BC, by simply putting the disclosure before the advice. This change does appear to reduce the laypersons’ errors by about \$0.24 in the experimental sample, but one cannot reliably extrapolate such findings since the estimate is far from statistically significant.¹¹⁰

Nonetheless, one might further hypothesize that the anchoring effect will be strongest during the layperson’s first estimation task, and that as he proceeds through the second through sixth estimation tasks (recall that there were six jars), he has internalized the information, and thus performs quite like those in the control group of 1BC. This dilution effect would not occur in one-off transactions, and thus the current intervention may still have policy relevance for such situations.

This new hypothesis can be tested by examining only the laypersons’ estimates on the first jar in the 1BC condition versus the first jar of the 1BCF condition. Indeed, when the disclosure is put first in 1BCF, layperson inaccuracy was improved by \$1.06 ($p = 0.04$).¹¹¹ The more precise hypothesis is thus confirmed, and this evidence suggests that disclosure policies should, where practicable, target laypersons *before* they receive substantive advice from conflicted advisors. Disclosures seem to work better as a prophylactic than as a remedy.

¹⁰⁸ Amos Tversky & Daniel Kahneman, *Judgment Under Uncertainty: Heuristics and Biases*, 185 *SCIENCE* 1124, 1128–30 (1974).

¹⁰⁹ Cain et al., *supra* note 32, at 6.

¹¹⁰ $M_{1BCF} = 6.25$ ($SE = 0.30$), $M_{1BC} = 6.49$ ($SE = 0.30$), $t(228) = -0.57$, $p = .57$. After initially finding a similar result, the researcher deployed conditions 1BC and 1BCF again in order to reduce the odds of incorrectly affirming the null hypothesis, thus resulting in double-sized samples. Even after these extraordinary efforts, the finding is far from significant.

¹¹¹ $M_{1BCF-Jar1} = 4.57$ ($SE = 0.28$), $M_{1BC-Jar1} = 5.63$ ($SE = 0.41$), $t(228) = -2.11$, $p = .04$, $r = .11$.

Nonetheless, this finding should be put in the context of condition 1BN, where there was one biased expert, with no disclosure given at all. For the first jars in 1BN, laypersons erred by \$4.28 on average, which is statistically indistinguishable compared to a disclosure-first policy (\$0.29 difference, $p = 0.6$).¹¹² Thus, putting disclosures first seemed to help ameliorate the problems with disclosure mandates in this experimental setting, but disclosure mandates were still worse than doing nothing about conflicting interests. An improved disclosure mandate thus appears to be a poor policy response to conflicts of interests, in this particular epistemic setting. Such a mandate seems to do nothing more than paper over a real problem for laypersons.

B. *What to Disclose*

Consider another method for improving the efficacy of disclosures. The CLM authors recognized that a disclosure of *conflicting interests* may not be particularly helpful to laypersons, because it does not provide information about whether the advisor is actually *biased* in her advice and, if so, to what degree.¹¹³ Indeed, in 1BC (as in the CLM study), laypersons were merely told, “Note: The advisor is paid based on how HIGH you are in estimating the worth of the jar of coins.”¹¹⁴ Laypersons were left to speculate about how these interests actually impacted the advice given. In principle, this need not be the case; at least in some contexts, policymakers could provide better information to laypersons. This could be a practicable policy solution in the information age, where massive datasets and statistical methods may allow a regulator to monitor the behaviors of conflicted versus non-conflicted advisors (whether physicians, stockbrokers, or mortgage brokers), with resolution at a group level or perhaps individual level. Indeed, pharmaceutical companies already use such “datamining” techniques to customize their marketing efforts to low-prescribing and high-prescribing doctors.¹¹⁵ Several states, and now the federal government, are developing databases of which physicians have relationships with pharmaceutical and device companies.¹¹⁶ If such behavioral information were collected by a regulator, paired with conflicts information, analyzed in a useful way, and passed along to the laypersons who rely upon

¹¹² $M_{1BN-Jar1} = 4.28$ ($SE = 0.50$), $M_{1BCF-Jar1} = 4.57$ ($SE = 0.28$), $t(155) = -0.53$, $p = .60$.

¹¹³ Cain et al., *supra* note 32, at 20–21.

¹¹⁴ *Id.* at 10.

¹¹⁵ Robert Post, *Prescribing Records and the First Amendment—New Hampshire’s Data-Mining Statute*, 360 *NEW ENG. J. MED.* 745, 745 (2009).

¹¹⁶ Weintraub, *supra* note 29.

conflicted advisors, it would thereby allow the layperson to more precisely discount the advice given.

Notably, such a policy mandating disclosures of bias may have different effects on the *advisors'* behavior than a policy that merely requires disclosure of conflicting interests. Advisors may not even know that they are biased by their conflicting interests.¹¹⁷ If advisors were simply told this information, then the social norming literature would suggest that the advisors might then change their behavior toward the norm.¹¹⁸ Imagine, for example, that hospitals in McAllen, Texas, and other extremely high-cost regions were required to disclose to their patients that, even controlling for population health, they charge more than twice as much per person as other hospitals, yet the quality of care and patient outcomes are statistically indistinguishable from that of other hospitals. One might suppose that this sort of mandate would cause the physicians and other advisors to improve their behavior, so as to reduce or eliminate the need for such embarrassing admissions in the future. In principle, this sort of intervention could completely ameliorate the advisor-side problem with disclosures. On the other hand, one might hypothesize that this data would simply provide advisors with even more "moral license" to give even more biased advice, as CLM observed with regular disclosures of conflicting interests.¹¹⁹ Perhaps this would be *caveat emptor* taken to the extreme. Resolving these competing hypotheses would be a fruitful avenue for future study. In any case, the present experiment does not measure the advisors' performance under this condition, but instead merely uses the advisor behavioral data from CLM's high disclosed condition, thus tacitly assuming that there would be no difference in advisor behavior.

Condition 1BCB (one biased advisor, with a disclosure of both the conflict and the average size of bias) tests this hypothesis, focusing just on how the strengthened disclosure would impact laypersons. In addition to a disclosure of conflicting interests (as in 1BC), condition 1BCB provides laypersons with more concrete information about the size of the conflicted expert's *bias* (rather than merely his conflicted interests). Specifically, in this condition, the

¹¹⁷ Moore et al., *supra* note 16, at 11 ("We argue . . . that doctors' advice is biased . . . and that they typically *believe* their biased advice is unbiased."); *see also* Gawande, *supra* note 1, at 40 (discussing how health care providers with a bias toward high-cost procedures treat patients without realizing the bias).

¹¹⁸ *See* Cass R. Sunstein, *Social Norms and Social Roles*, 96 COLUM. L. REV. 903, 930, 949 (1996) (considering how choices are based upon beliefs about facts, and how the communication of accurate facts can therefore change beliefs based on inaccurate facts).

¹¹⁹ *See* discussion *supra* Part II.

experimenters told the subjects that “prior research has shown that advisors paid in this way tend to give advice that is \$7.68 higher on average than the advice of advisors who are paid based on accuracy.” This was a true statement, based on the data reported in CLM¹²⁰ and the prompts used in the present experiment.

Compare condition 1BCB against condition 1BC on the dependent variable of layperson accuracy. The addition of an average bias disclosure did not help layperson accuracy on average (but may have actually worsened it by \$0.46 on average, although this is statistically insignificant, $p = 0.47$).¹²¹ The hypothesis is rejected—a disclosure of the conflicted advisors’ average level of actual bias does not appear to help the average accuracy of laypersons that rely upon those advisors. Another condition, 1BCBF, further suggests that it makes little difference when this bias information is disclosed, whether first, before the substantive advice, or thereafter. Like condition 1BCB, condition 1BCBF provided laypersons with disclosures about average advisor bias, but did so first, before providing the advisor’s substantive advice. The slight improvement of \$0.21 over 1BCB is not significant ($p = 0.81$).¹²² Thus, the hypothesis that disclosing actual bias will help laypersons discount optimally must be rejected.

Although laypersons could have simply subtracted \$7.68 from the advice they received, and thereby calculated (and used) the same advice received by laypersons with unbiased advisors (on average), they apparently did not do so. Why did this intervention fail? Participants were allowed to answer an optional final question, providing open-ended feedback on the study or describing their tactics, and some of the answers are relevant to this point. Although a few participants said, “I pretty much just subtracted the \$7.00,” as one would hope and expect, others receiving this bias disclosure said, “I pretty much ignored the adviser, they seemed like they were way off, and knowing they were biased meant there was no reason to take their word.”¹²³ A significant number of respondents used the bias disclosure not as a mechanism

¹²⁰ See Cain et al., *supra* note 32, at 15 tbl.5.

¹²¹ $M_{1BCB} = 6.95$ ($SE = 0.64$), $M_{1BC} = 6.49$ ($SE = 0.30$), $t(170) = 0.73$, $p = .47$.

¹²² $M_{1BCBF} = 6.74$ ($SE = 0.55$), $M_{1BCB} = 6.95$ ($SE = 0.64$), $t(120) = -.247$, $p = .81$.

¹²³ A third subgroup of respondents in 1BCB and 1BCBF seems to have actually been misled by the disclosure of bias and provided even higher raw guesses than in the 1BC condition, drawing the average guess higher. Other than sheer confusion, or a failure to communicate clearly, no obvious hypothesis explains why this might happen. The increased standard deviation that comes with an average bias disclosure (3.28 in 1BC to 4.79 in 1BCB) suggests that there is more than simply a shift in means occurring in this data.

of calibrating their reliance more precisely, but rather as a strengthened warning suggesting that the advice is altogether worthless. Given the high levels of epistemic asymmetry in this experiment (measured in the prior Part), the tactic of ignoring the proffered advice turns out to be a very poor idea. Still, the findings in the prior section suggest that a specific bias disclosure may be more fruitful in contexts of low epistemic asymmetry (such as stock broker–client relationships or realtor–buyer relationships), as it would drive laypersons away from advice that had very little value in the first place. This deserves further study, in various epistemic contexts.

To disaggregate these trends, let us create a benchmark for layperson success in this task. Suppose that condition 1UA presents a decent benchmark for success, since it provides laypersons with one unbiased advisor and a statement that interests are aligned. The researcher constructed a proportional metric representing the percentage of participants in each condition whose guesses were as good or better than the \$2.72 benchmark error of a median respondent in the 1UA condition. Let us stipulate that the participants more or less “succeeded” in the estimation task, if their inaccuracy was no worse than the laypersons’ in the 1UA condition. By definition, 50% of the participants in 1UA performed at or better than their own median, but when the disclosed conflict is added in 1BC, only 11% exceeded the benchmark for success. However, when a policymaker added a mandate for disclosure of actual average bias in 1BCB, the “successes” increased to 21%. We have nearly doubled the number of successes.¹²⁴ As a matter of public policy, this could be a worthwhile investment, if it doubled successes, helping laypersons to overcome a given threshold and make better decisions (e.g., rejecting the gallbladder surgery recommended by their conflicted surgeon where there is no proven marginal efficacy).

The conclusions here are very tentative. It may be worthwhile to further pursue the concept of mandating disclosures of biased advisor behavior, perhaps with special attention to making the information useable to the laypersons who must rely upon it, so as to minimize confusion and maximize their ability to integrate the additional information into their process of weighing the advice against their own epistemic priors. At the end of the day,

¹²⁴ This result is marginally significant at traditional levels. Using a chi-squared test comparing 1BC with 1BCB, $\chi^2(1) = 3.18, p = .06$. When 1BCB is combined with the statistically indistinguishable 1BCBF (where the only difference is that the disclosures are provided first, before the advice) and then compared with 1BC, the difference in success rates is significant, $\chi^2(1) = 3.82, p = .04$; the odds of “success” were 1.98 times higher in the (combined) 1BCB+1BCBF condition than in the 1BC condition.

this intervention may have a distributive effect, helping the savviest laypersons weigh the information they receive, but harming others who react poorly to the additional information. These effects may depend in part on the degree of epistemic asymmetry (i.e., relative expertise) in a given context. As discussed above, only in situations of high epistemic asymmetry will it be worrisome for a policy to drive a wedge between a layperson and her advisor. And, as discussed further below, in a robust marketplace for advice, a disclosure of bias may have the salutary effect of driving laypersons to better advisors—a choice that laypersons did not have in the present experimental conditions. The concept of bias disclosures (rather than conflict disclosures) thus deserves further study in other experimental and policy settings.

C. *To Whom to Disclose*

Consider a third potential way to improve disclosure mandates: by tailoring them to individual persons who need them while withholding them from laypersons who could only be harmed by them. Consider the likely real-world contexts in which a conflicting interest exists but some advisors remain unbiased—they do not change the advice that they give to some or all of their layperson clients, compared to the advice they would have given but for the conflict. Heterogeneity arises at two levels: (1) that of the individual advisors and (2) that of the individual laypersons who rely upon them.

First, advisors' professionalism—their technical training and ethical commitments—may prevent some of them from suffering biases, even when they have conflicting interests.¹²⁵ Even if the mean advice differs between conflicted advisors and non-conflicted advisors (as CLM reported, and we assume here), the distributions of the two groups are likely to overlap, such that a significant portion of the conflicted advisors will perform as well or better than the median non-conflicted advisor. The mere fact that an expert is conflicted does not necessarily imply that his advice is biased.¹²⁶

The phenomenon repeats at the level of the individual layperson clients within each advisor. Even within the biased advisors, only some of their

¹²⁵ See Robertson, *supra* note 47, at 193–95 (discussing the ways in which professionalism constrains the biases of experts, albeit imperfectly).

¹²⁶ See, e.g., *Pretty v. Prudential Ins. Co. of Am.*, 696 F. Supp. 2d 170, 189 (D. Conn. 2010) (“The mere fact that Prudential retained the medical experts to review the Plaintiff’s file does not make their opinions unreasonable. The Plaintiff has also failed to provide any evidence of a history of biased claims administration by Prudential.” (citation omitted)).

clients will receive biased advice compared to what they would have received from an unbiased advisor. This ratio will be particularly low where the advisor provides a binary sort of advice, as is often the case. For example, a doctor may advise either treatment *S* (surgery) or treatment *L* (lifestyle changes). Even if such an advisor becomes biased, this will just increase the frequency with which he gives the favored advice (*S*). Without the conflicting interest, a given doctor may have prescribed the surgery to 70% of his clients presenting with a given condition, but after succumbing to the bias, he then prescribes it to 85% of his clients. For most of the clients (aside from the marginal 15%), the substantive advice will be the same in either case, but the advice will now be accompanied by a warning about conflicting interests.

To simulate the performance of that majority group, participants in condition 1UC each received one unbiased advisor (as in 1UA) but a disclosure of conflicting interests (as in 1BC). As one might hypothesize, these laypersons suffered from the disclosure, having errors \$1.21 larger on average than those in condition 1UA ($p = 0.049$).¹²⁷ Thus, these findings illustrate how, in the real-world settings of doctors' offices and mortgage brokerages, a disclosure mandate may often drive laypersons away from perfectly good advice. This is an important finding, identifying and demonstrating another way that disclosures may be deleterious to the people that they are designed to help.

From the perspective of layperson welfare, this is another piece of evidence that suggests that disclosure mandates are poor solutions for the problem of conflicting interests. The real solution would try to eliminate the conflicts in the first place. Still, if we continue to rely on disclosure mandates at all, as seems inevitable, it may then be best to narrow disclosure mandates to only those situations where we have some reason to believe that a particular advisor or set of advisors is actually biased (not merely exposed to a *potential* bias arising from a conflicted interest). Even better, we would further limit disclosure to those particular laypersons who are receiving the marginal advice that is different from what would have been given but for the bias. For example, as discussed above, there are extreme geographic disparities in health care costs across the United States, with health care providers in some regions charging for twice as many procedures compared to others, with no discernable improvement in quality.¹²⁸ In principle, a disclosure mandate could target only

¹²⁷ $M_{1UC} = 4.77$ ($SE = 0.42$), $M_{1UA} = 3.56$ ($SE = 0.42$), $t(90) = 2.00$, $p = .049$, $r = .21$.

¹²⁸ Orszag & Ellis, *supra* note 5, at 1794–95.

the regions or institutions where costs are highest, where regulators expect that it is most likely that patients are suffering from biased advice. Thus, any benefits of a disclosure mandate can be captured without imposing the costs identified here. Or more particularly, depending on the resolution of the data, the mandate could be tailored to individual hospitals or even individual doctors.

In principle, targeted disclosures can work at the patient level. Scholars have found that doctors tend to practice quite similarly when the evidence and national guidelines are clear, but in some regions they exhibit biases for higher cost care when they make decisions under greater uncertainty.¹²⁹ Thus, to the extent that such situations can be identified *ex ante*, a disclosure mandate could be required for those situations but not others. As Margaret Johns has proposed, regulators could require physicians to disclose conflicts of interest when they write off-label prescriptions, but the regulators need not require disclosures when conflicted doctors prescribe on-label or in accordance with national practice guidelines.¹³⁰

Putting aside this possibility of narrowly tailored disclosure mandates, the bottom-line finding of condition 1UC is important to emphasize. For another reason, crude disclosure mandates can be deleterious to the laypersons they are designed to help. Putting autonomy-based arguments aside, policymakers concerned with patient welfare should be careful not to force disclosures of conflicting interests unless they have credible evidence that the conflict actually causes a bias for the layperson, and evidence that the disclosure will make things better.¹³¹ Furthermore, if they have such evidence of actual bias, the disclosure mandate should be tailored as narrowly as possible to specific groups of advisors and laypersons. Then, as discussed in Part III.B, the

¹²⁹ Brenda E. Sirovich et al., *Regional Variations in Health Care Intensity and Physician Perceptions of Quality of Care*, 144 ANNALS INTERNAL MED. 641, 646 n.2, 648 (2006) (examining how doctors with poor communication with patients, restrictions upon autonomy, and a perceived scarcity of resources result in a higher cost of care).

¹³⁰ Johns, *supra* note 26, at 971. The FDA apparently prohibits physicians with industry ties from promoting a drug for an off-label use but allows industry-tied physicians to prescribe a drug for off-label use. See Conko, *supra* note 37, at 15.

¹³¹ A fair question arises about the default rule. It may be a decent assumption that wherever there is a conflict of interest there is probably a bias in the aggregate advice rendered. The argument of this section has merely sought to show that there is a heterogeneity of advisors and a heterogeneity of laypersons, such that a statement about the aggregate cannot reliably be applied to each piece of advice individually. Such generalization would be an example of the ecological fallacy. See generally GARY KING, A SOLUTION TO THE ECOLOGICAL INFERENCE PROBLEM 3–17 (1997) (discussing ecological inferences and the ecological fallacy).

evidence of bias should perhaps be provided to laypersons themselves so that they can better assess the advice that they receive.

IV. CALIBRATING RELIANCE IN A MARKET FOR ADVICE

Part III explored ways to improve the efficacy of disclosure mandates. Even with such improvements, however, disclosures are likely to remain a suboptimal, or at least incomplete, solution for the fundamental problem of biased advice. One remaining hypothesis, not tested by CLM or the foregoing experimental conditions, is that disclosures may help laypersons choose amongst multiple conflicted and non-conflicted advisors if there is something like a market for advice. This Part applies several new experimental conditions to explore laypersons' baseline assumptions about advice, and whether affirmative disclosures may improve reliance and performance when interests are aligned. This Part also introduces several conditions in which laypersons are given multiple biased and unbiased advisors, with and without conflicting interests. Finally, by assessing the correlation between layperson confidence and performance, this Part concludes that market-based solutions are likely insufficient. Laypersons appear to have little self-awareness about their marginal performance with or without biased advisors, which thus makes more aggressive regulatory interventions appropriate.

A. *Affirmative Disclosures of Aligned Interests*

Almost two decades ago, scholars in biomedical ethics were already identifying a crisis in trust—patients had reduced their degree of reliance on their health care providers, to the detriment of both the patients' health outcomes and the esteem of the medical profession.¹³² In dentistry, for example, the fee-for-service relationship creates deep conflicting interests, and there is even less oversight by insurers and government payors.¹³³ Dentists have begun to worry about polling data showing that the U.S. public trusts their honesty and ethics at a rate lower than that of nurses, pharmacists, and physicians.¹³⁴ The longer dentists have practiced, the more they are conscious

¹³² Edmund D. Pellegrino, *Trust and Distrust in Professional Ethics*, in *ETHICS, TRUST, AND THE PROFESSIONS* 69, 77–78 (Edmund D. Pellegrino et al. eds., 1991).

¹³³ See, e.g., *United States v. Talbott*, 590 F.2d 192, 195–96 (6th Cir. 1978) (upholding rare convictions for mail fraud for unnecessary dental procedures).

¹³⁴ Barry Schwartz et al., *Perceptions About Conflicts of Interest: An Ontario Survey of Dentists' Opinions*, 71 *J. DENTAL EDUC.* 1540, 1548 (2007).

of the problems created by their conflicting interests.¹³⁵ Such lack of trust may mean that skeptical patients forego needed dental work.

Some scholars have suggested that disclosure policies may be part of the solution to this problem of diminishing trust in professional advisors.¹³⁶ Kevin Weinfurt, for example, hypothesized that in contexts of high epistemic asymmetry (as here), where a layperson does have an advisor whose interests are aligned, a disclosure of that fact may help the layperson become more accurate by making the layperson more trusting.¹³⁷ Condition 1UN of this study, which had one unbiased advisor but no such disclosure, was designed to test this hypothesis against condition 1UA, where there was also one unbiased advisor and laypersons were told, "Note: The advisor is paid based on how accurate the estimator is in estimating the worth of the jar of coins," as in CLM's accurate condition.¹³⁸

The results were positive, showing that such an affirmative disclosure of aligned interests in 1UA improves layperson performance by \$1.15 on average compared to the agnostic 1UN ($p = 0.05$).¹³⁹ This finding demonstrates that in our experimental setting at least, laypersons were naturally rather untrusting of the advice that came with epistemic advantages but without any information about incentives. The information about the advisors' aligned incentives seemed to overcome this natural distrust and increased reliance accordingly.

This condition also allows us to isolate the effect of a disclosure of conflicting interests, while holding the substantive advice constant. Let us construct a measure of the layperson's degree of reliance on the expert's advice, defined as the difference between the advice given and the estimate rendered. The larger that difference, the less the layperson appears to be relying upon the expert. In 1UA (where there was an unbiased advisor and a disclosure of aligned interests), laypersons' estimates were on average \$3.64 away from the advice given, while those receiving a conflicts disclosure in 1UC were on average \$5.15 away from the advice given, a difference of \$1.51 ($p = 0.02$).¹⁴⁰ In 1UN (where the unbiased advice was the same but there was

¹³⁵ *Id.* at 1548.

¹³⁶ Kevin P. Weinfurt et al., *Disclosing Conflicts of Interest in Clinical Research: Views of Institutional Review Boards, Conflict of Interest Committees, and Investigators*, 34 J.L. MED. & ETHICS 581, 581, 585 (2006).

¹³⁷ *Id.* at 581–83.

¹³⁸ Cain et al., *supra* note 32, at 10.

¹³⁹ $M_{1UN} = 4.71$ ($SE = 0.38$), $M_{1UA} = 3.56$ ($SE = 0.42$), $t(89) = 2.00$, $p = .049$, $r = .21$.

¹⁴⁰ $M_{1UA} = 3.64$ ($SE = 0.46$), $M_{1UC} = 5.15$ ($SE = 0.45$), $t(90) = -2.30$, $p = .02$, $r = .24$.

no information about the advisor's incentives provided), the layperson on average provided estimates that were \$4.94 away from the advice given. Thus, when no incentives information is provided, as in 1UN, laypersons seem to behave almost exactly the same as when a conflict is disclosed, as in 1UC (a difference of \$0.21, $p = 0.74$).¹⁴¹ This is quite surprising, given that the experiment provided no prompting at all that would suggest that the advisor may have a conflicting interest or any motives whatsoever other than truth. Nonetheless, the disclosure that interests were aligned in 1UA improved reliance and accuracy significantly.

Thus, in real-world settings where advisors and clients have aligned interests, a disclosure mandate may help laypersons properly increase their reliance. Of course, if greater reliance is in the advisor's own interests, a mandate may be unnecessary. However, it is also possible that social norms or sheer habit will prevent overt discussion of the advisor's incentives. This failure is especially likely where the policy regime has not yet focused attention on those incentive structures. Thus a disclosure mandate policy, designed to help laypersons with conflicted advisors, may have spillover benefits to even those with non-conflicted advisors. This finding also holds promise for policies that explicitly attempt to align the incentives of advisors and laypersons, suggesting that laypersons would be quite appreciative of such reforms and that their behavior would exploit such an improvement, if they learned about it.

B. Using Disclosures to Select Advisors

A significant limitation of the CLM study was that laypersons receiving disclosures about conflicts had nowhere else to turn for advice. Each layperson had a single advisor, who essentially had a monopoly on the market for advice. If the layperson did not trust her advisor's opinion, she could only resort to her own inferior estimates. Instead, as CLM acknowledged but did not test,¹⁴² one might hypothesize that a disclosure mandate will be salutary to laypersons when there are multiple biased and unbiased advisors available because it helps the laypersons decide where to place their reliance. Indeed, this selection effect may be the most important function of a disclosure mandate in real-world settings.

¹⁴¹ $M_{1UN} = 4.94$ ($SE = 0.44$), $M_{1UC} = 5.15$ ($SE = 0.45$), $t(103) = -0.33$, $p = .74$.

¹⁴² Cain et al., *supra* note 32, at 21–22.

This logic seems to be the assumption behind rules that allow an advisor to proceed with a conflicting interest, as long as that interest is first disclosed to the client and the client gives informed consent.¹⁴³ As formal models predict, this layperson-choice dynamic might also then create an incentive for the expert to either eliminate the conflicting interests or credibly demonstrate to the potential layperson clients that he is nonetheless unbiased.¹⁴⁴ On the other hand, in specific contexts, such as the conflict created by attorney referral fees, scholars have argued that a ban may be more efficient than a disclosure mandate.¹⁴⁵ States are experimenting with both approaches,¹⁴⁶ but there would seem to be little means of assessing the success of these natural experiments.

The present laboratory experiment does not test the *ex ante* effects on advisors, but it does test the possibility of *ex post* benefits of disclosure mandates to laypersons in multi-advisor settings. Conditions 2N and 2D each have two advisors per layperson, one of whom is biased by a conflicting interest. In 2D, but not 2N, a disclosure mandate is imposed, which worsens the advice of that advisor (as in CLM) but provides valuable information to the layperson. Thus, we have something like two miniature markets for advice, with a variety of both conflicted and non-conflicted advisors, and in one setting there is a disclosure mandate. Still, it is notable that, unlike a true market, the second opinion was automatically provided without imposing the cost thereof on the layperson.

Here, unlike in CLM's single-advisor experiment, the disclosure mandate has significant salutary results for the laypersons, improving accuracy by \$1.22 ($p = 0.02$).¹⁴⁷ On net, the advice received in the disclosure condition was worse than in the undisclosed condition, but the laypersons did not blindly average them. Apparently, the laypersons effectively used the conflicted

¹⁴³ See, e.g., MODEL RULES OF PROF'L CONDUCT R. 1.7 cmts. 18–19 (2010).

¹⁴⁴ Joel Sobel, *A Theory of Credibility*, 52 REV. ECON. STUD. 557, 557–58, 570 (1985).

¹⁴⁵ John S. Dzienkowski & Robert J. Peroni, *Conflicts of Interest in Lawyer Referral Arrangements with Nonlawyer Professionals*, 21 GEO. J. LEGAL ETHICS 197, 235 n.181 (2008).

¹⁴⁶ *Id.* at 208 n.71, 210 n.79.

¹⁴⁷ $M_{2N} = 3.97$ ($SE = 0.42$), $M_{2D} = 2.75$ ($SE = 0.27$), $t(108) = 2.43$, $p = .02$, $r = .23$. As will be explained further below, there is also an effect for simply providing two advisors rather than one. If you compare 1BC against 2D—that is, one versus two advisors, with disclosures in both—estimators with two advisors do much better, $M_{2D} = 2.74$ ($SE = 0.27$), $M_{1BC} = 6.49$ ($SE = 0.30$), $t(98) = 1.47$, $p < .001$, $r = .15$. Interestingly though, there is no significant difference between using one versus two advisors if there is no disclosure (that is, comparing 1BN against 2N), $M_{2N} = 3.97$ ($SE = 0.42$), $M_{1BN} = 2.65$ ($SE = 0.40$), $t(98) = 1.47$, $p = .14$. In short then, with one conflicted advisor, laypersons do better with no disclosure, but with one conflicted and one non-conflicted advisor, they do better with disclosure. This interaction reaffirms the point that a disclosure is only useful to laypersons if laypersons have somewhere else to turn for advice.

advisors' disclosures to place their reliance on the non-conflicted advisors. Thus, when laypersons have access to non-conflicted advisors, a disclosure may be salutary.

C. *The Value of Second Opinions*

For policymakers then, a primary challenge is to get non-conflicted advisors to the laypersons who need them. One such mechanism is exemplified by regulations that mandate that patients or mortgage borrowers get second opinions before acting on the advice of conflicted advisors.¹⁴⁸ Is a second-opinion an antidote to biased advice, or are more radical remedies (such as a ban on the biased advice) necessary?

Comparing condition 2D with 1BC allows one to test such "second-opinion" policies. Laypersons receive biased advice with a disclosed conflict in each condition, but in 2D, the layperson also receives a "second opinion" from an unbiased advisor. This intervention dramatically improves layperson performance by 53% (a difference of \$3.47, $p < 0.001$).¹⁴⁹ This is one of the starkest differences in layperson performance reported in this study. Indeed, this 2D condition becomes the new gold standard for layperson accuracy, marginally improving on even condition 1UA (CLM's "accurate"), where a single advisor has aligned incentives for accuracy (a difference by \$0.81, nearly significant, $p = 0.09$).¹⁵⁰

One might worry that second-opinion policies will be limited in their effectiveness if the layperson anchors on the first advice received and does not sufficiently adjust his estimate upon receiving the new advice.¹⁵¹ Further experimental research, described in the footnotes, explores and rejects this hypothesis.¹⁵²

¹⁴⁸ See *supra* notes 39–41 and accompanying text.

¹⁴⁹ $M_{2D} = 2.74$ ($SE = 0.27$), $M_{1BC} = 6.49$ ($SE = 0.30$), $t(98) = 1.47$, $p < .001$, $r = .15$.

¹⁵⁰ $M_{1UA} = 3.56$ ($SE = 0.42$), $M_{2D} = 2.75$ ($SE = 0.27$), $t(90) = 1.70$, $p = .09$, $r = .18$. How can adding a biased-disclosed advisor (as in 2D) improve performance over simply receiving advice from an advisor with aligned incentives (as in 1UA)? It may be that the two pieces of advice were relatively coherent, compared to the layperson's own estimate (which we know from condition NoAdvisors is much further from the truth). Thus the biased, disclosed opinion communicated a rough scale of the epistemic asymmetry to the layperson, helping her to place closer reliance on the unbiased expert's advice. This effect may be peculiar to settings of high epistemic asymmetry and relatively low bias ratios.

¹⁵¹ See Tversky & Kahneman, *supra* note 108.

¹⁵² Condition 2DR was designed to test the hypothesis that laypersons would anchor on the first advice received. Condition 2DR is identical to condition 2D, except that the order of advisors is reversed, so that the layperson first receives advice from an advisor with aligned incentives (and a disclosure of the same) and then

Notably, the experiment assumed that the biased advisor would perform the same as he would under a condition of mandatory disclosure alone. In the real world, if a biased advisor knows that the layperson would likely receive a second opinion from an unbiased advisor, the biased advisor may perform differently, perhaps improving the advice she gave, making the net effectiveness of this policy even better. Future studies should test this potential improvement in *advisor* behavior.

The most important policy-relevant conclusion remains: the clearest remedy to the epistemic asymmetry with conflicting interests is to (a) force disclosures of the conflicts, but only if we also (b) ensure that laypersons also have access to, and actually use, non-conflicted advisors. Non-conflicted advisors are a complete antidote to conflicted advisors. Of course, second opinions have costs—someone must pay that second advisor to repeat the work of the first conflicted advisor. The present experiment did not impose those costs on the laypersons, but instead provided the second opinion for free. It is a context-dependent empirical question whether the additional costs will outweigh the biases imposed by the first conflicted advisor. Thus, policies that align incentives in the first place may be more efficient, if there are viable mechanisms for such alignment in a given market.

D. A Market for Unbiased Advice

The foregoing findings suggest that, to some extent, laypersons can themselves incentivize the production of unbiased advice, since in a regime of disclosure the laypersons tend to follow non-conflicted advice over conflicted advice. Therefore, if competition exists amongst advisors in something like a market, those who credibly avoid conflicting interests may demand a premium for their advice.

receives advice from an advisor with conflicting interests (and a disclosure of the same). Surprisingly, laypersons in 2DR perform significantly worse than those in 2D, where the accurate advice is second (a \$0.81 difference), $M_{2D} = 2.74$ ($SE = 0.27$), $M_{2DR} = 3.56$ ($SE = 0.29$), $t(103) = -2.04$, $p = .04$, $r = .20$. Thus, sequencing does seem to matter. A second opinion is apparently more influential, either because the first opinion's disclosure of conflicts primes the layperson to look for a more reliable source of advice, which then becomes particularly compelling once found, or perhaps simply because the second source of advice is more proximate in time to the decision task, which immediately follows. Nonetheless, even with this suboptimal sequencing and the inclusion of a biased advisor, the point estimate for layperson inaccuracy in 2DR is \$3.56, precisely the same as in the previous gold-standard condition of 1UA (a single advisor with aligned incentives).

In some contexts, there may be market actors that will benefit from laypersons following unbiased advice and have a mechanism for providing such unbiased advice. If a producer truly does sell the better product, it will prefer that laypersons get non-conflicted advice that will help them choose the better product. That is why, for example, carmakers like to brag that a purportedly independent expert, such as J.D. Power and Associates, provides favorable advice.¹⁵³ Likewise, in litigation, attorneys often prefer to use treating physicians as expert witnesses, since they render opinions without influence by the lawyers, in contrast to hired-gun experts that were handpicked and coached by attorneys.¹⁵⁴ A law that mandates disclosure of conflicts may help create such a market for advice, if it draws laypersons' attention toward this issue, and if there are unbiased sources of advice available.

Nonetheless, a market mechanism would require laypersons to know what sort of advice they need and to be willing to pay for it. It bears emphasis that the experimental conditions with multiple advisors did not require the laypersons to realize that they needed a second opinion, or to pay for it. The second opinion just appeared alongside the first. In the real world, advice will always have a cost, at least in terms of time and inconvenience, if not in service fees charged by the advisor. Further, unbiased advice will tend to be more expensive to a layperson than biased advice, which a third party subsidizes.

Do laypersons have the necessary meta-knowledge, i.e., an understanding of their own epistemic strengths and weaknesses, in advisory situations? Do they know whether they need advice, and if so, whether they need non-conflicted advice, and at what price? The laboratory experiment allows us to approach these questions, albeit only indirectly. After seeing the photograph of each jar, receiving the advice and disclosures (if any), and rendering their own estimates, participants were asked, "How confident are in your estimate? (10 = very confident, 1 = not confident)."¹⁵⁵ The CLM study did not provide

¹⁵³ See, e.g., Cambridge PR Group, *Ford Surges in J.D. Power and Associates Initial Quality Survey*, READMEDIA, July 13, 2010, <http://readme.readmedia.com/Ford-Surges-in-J-D-Power-and-Associates-Initial-Quality-Survey/1591858> (touting Ford Motor Co.'s "huge accomplishment" of ranking favorably in the J.D. Power and Associates survey).

¹⁵⁴ See Robertson, *supra* note 47, at 194–95. Indeed, litigants could use a blinding mechanism to more regularly bring unbiased experts to trials as a rational strategy for garnering extra credibility from the lay fact-finder. See *id.* at 215.

¹⁵⁵ This relationship between subjects' confidence and accuracy in an estimation task, which is known as calibration, has been extensively studied in the judgment and decision-making literature. See generally Claire

such a measure of layperson self-assessment, but it may be useful as a proxy for how laypersons will perform in a market for advice. When given incentives for accuracy, one might expect that laypersons would be willing to pay more to move to positions of higher confidence as their own best proxy for accuracy. For this proxy to be effective, and for the market to work, there must be a significant correlation between the accuracy of layperson estimates and their confidence in their estimates.

Such a hypothesized relationship is not apparent in this data. Across all conditions of the study, there is no relationship between the average accuracy of laypersons' judgments and their average expressed confidence in those judgments.¹⁵⁶ The participants apparently had no idea as to whether they were doing well or poorly. In contrast, one would have hoped that those in the inaccurate conditions, such as those with no advisor or an advisor with a disclosed conflict of interest, would express low confidence, such that they might be willing to pay a premium to move to a more accurate condition. This was not the case.

It gets worse when the participants are clustered into the twelve experimental conditions, as shown in Figure 1. There was significant variation in average confidence levels between conditions, ranging from 4.69 for condition 2D to 6.07 for condition NoAdvisors, a difference of 1.38 on the 10-point Likert scale.¹⁵⁷ Notably, the participants in the NoAdvisors condition are much more confident than participants in any other condition, even though they perform much worse than in any other condition. Indeed, there is a strong correlation between average layperson *inaccuracy* by condition and average confidence by condition.¹⁵⁸ To be clear, this would be a negative correlation between *accuracy* and confidence.

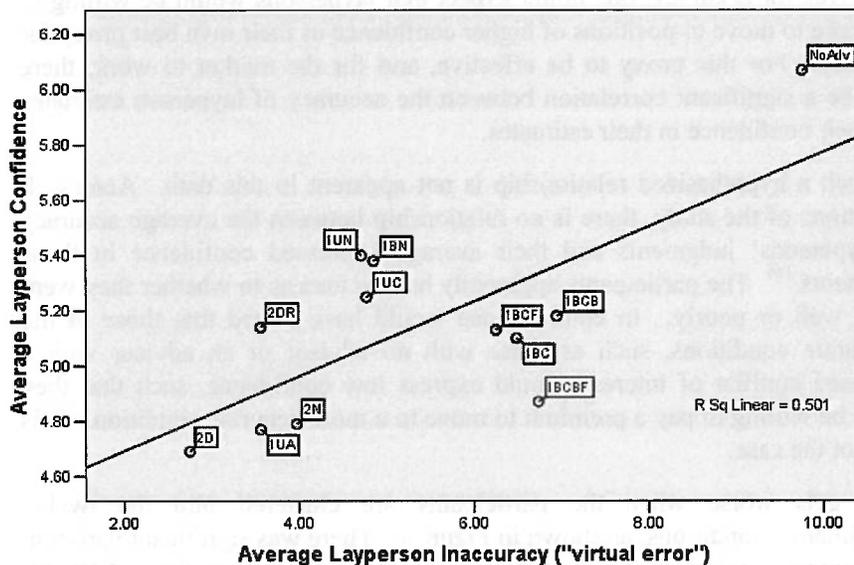
I. Tsai et al., *Effects of Amount of Information on Judgment Accuracy and Confidence*, 107 ORGANIZATIONAL BEHAV. & HUM. DECISION PROCESSES 97 (2008) (reviewing this literature).

¹⁵⁶ Pearson $r < 0.01$, p (two-tailed) = .95.

¹⁵⁷ $F(2, 611) = 1.99$, $p = .04$, $r = .17$; see *infra* Table 4.

¹⁵⁸ Pearson $r = .71$, p (two-tailed) = .01.

Figure 1:
Layperson Inaccuracy and Confidence by Condition



This finding may be peculiar to the particular estimation task utilized in this study and in the CLM experiment. Since coins are a feature of daily life, laypersons may have a very high degree of confidence in their own abilities to render an accurate assessment, but they actually tend to systematically underestimate the value of coins. In the NoAdvisors condition, the laypersons may have been most confident because they received no information that would undermine their prior beliefs. In other conditions, when advice came from advisors with aligned interests (1UA) or from multiple advisors (2N and 2D), this advice was very persuasive to the laypersons, but it apparently created cognitive dissonance with the laypersons' prior beliefs and may have thereby undermined the laypersons' confidence. When the advice was even worse in the 1BC and 1BCB conditions, the laypersons could confidently disregard it as unreliable and proceed with their own estimates.

These findings illustrate the complexity of setting policy to improve laypersons' epistemic performance. One cannot blindly assume that

laypersons will pay for the quality of advice they need, or be able to assess accurately the quality of the advice that they receive.¹⁵⁹ Economists use the term *credence goods* for products, like expert advice, for which the buyer has little ability to monitor quality.¹⁶⁰ Unlike other products and services, the market for advice is defined by the layperson's own epistemic incompetence. This is especially true in the health care context. As Marc Rodwin explains, "[P]atients are particularly vulnerable. . . . They often have little opportunity to learn from personal experience, or the cost of doing so may be high. These constraints distort their choices as consumers and increase their reliance on the recommendations of their physicians."¹⁶¹ In other real-world settings, laypersons do receive feedback about the decisions they make—for example, they watch as their 401(k) accounts soar or fall compared to benchmark indexes. The present study provided laypersons with no such feedback. Still, in the real world, the feedback may come too late to be actionable and may come in forms that are not particularly intelligible to the layperson, if there are no clear comparisons or baselines available.

These experimental conditions suggest that much will depend on which advisors happen to get to laypersons first, because once a layperson–advisor relationship is created, it is likely to be sticky. A layperson with a highly conflicted advisor would appear to proceed with a high degree of confidence and would be unlikely to switch advisors. Unfortunately, in a market setting, the most highly conflicted advisors likely have the greatest incentives to aggressively find and recruit layperson clients. At this point, policymakers may have few options. Once the layperson has a relationship with a conflicted advisor, a mere disclosure may not be enough to break that connection. This is especially true when the costs of the conflicted advisor are completely subsidized or already sunk, but the layperson would have to pay for a second opinion from an unbiased advisor.¹⁶² Thus, the key task for policymakers is to find ways to get unbiased advisors to laypersons in the first place.

¹⁵⁹ See Hadfield et al., *supra* note 15, at 144 ("The complex nature of information also requires careful analysis of the potential for market mechanisms to provide the information consumers might want and need. Information is a notoriously difficult commodity over which to contract. Potential buyers of information have difficulty determining, in their uninformed state, the value of the information and thus the price they are willing to pay for it. Sellers of information run the risk of revealing their information, and thus the commodity they hope to sell, by the very terms on which they offer to sell. . . . These observations counsel care in relying on market information intermediaries to resolve the problems of information in consumer markets.")

¹⁶⁰ See Winand Emons, *Credence Goods and Fraudulent Experts*, 28 RAND J. ECON. 107, 107 (1997).

¹⁶¹ Rodwin, *supra* note 31, at 1406.

¹⁶² See Hadfield et al., *supra* note 15, at 145 ("Information is costly and so consumers rationally make choices between being better informed and settling for a less informed but less (transaction) costly option.")

CONCLUSIONS—ELIMINATING BIASES WITH SOUND POLICY

In the modern capitalist society, reliance relationships based on epistemic asymmetry will only grow in importance as transactions become more sophisticated and the need for specialization grows. It seems clear that conflicts of interests and resulting biases will only proliferate as those with expertise, or the appearance thereof, seek to exploit those advantages. Thus, this Article has sought solutions.

Still, this study had several noteworthy limitations. First, the coins-in-jars estimation task may not be comparable to all (or any) real-world contexts faced by laypersons. Future studies should create more realistic decision situations, such as that facing a patient deciding whether to take a prescribed drug as his conflicted doctor recommends, or that of an investor deciding whether to buy the stock recommended by the conflicted advisor. The advantage of the coins-in-jars estimation task is that it was concrete (with a right or wrong answer knowable by the researcher), and it was conducted realistically (human subjects were not asked to pretend that they were actually a patient in a treatment situation). This study also lacked feedback for laypersons, which may be present in some real-world situations. Future studies should also employ a more nuanced model of the market for multiple sources of advice, allowing laypersons to choose whether to purchase second opinions, and impose transaction costs on those choices. Further studies should also explore the effects on *advisors* of the various policy mechanisms tested on laypersons here, including actual bias disclosures and second opinion mandates.

The present empirical study has yielded several important conclusions. First, it has added further credence to Cain, Loewenstein, and Moore's observation that disclosure mandates can make matters worse, if they worsen the advice given but fail to help laypersons truly improve their own estimates. By measuring epistemic asymmetry (relative expertise) and the degree of advisor bias, however, the present study has revealed the contingent nature of such conclusions. This more nuanced account allows analysts to begin thinking more clearly about the contexts in which a disclosure mandate or even a ban on conflicted advice may be worthwhile. Although epistemic asymmetry is a pervasive feature of modern life, so too is epistemic charlatanism and biased advice. In these situations, a disclosure mandate may be salubrious, if it drives laypersons away from bad advice.

Still, the present study has explored several ways to improve disclosure mandates, even where expertise is real. For initial interactions with advisors, it

helped to provide disclosures before conflicted advice, but the effect diminished with iterative interactions with the same advisor.

The study also explored the possibility of implementing disclosure mandates that focus on actual biases, rather than mere conflicts of interest. The present study found, however, that disclosure of actual advisor bias did not improve average performance compared to disclosure of mere conflicting interests. Still, bias disclosures did help significant portions of the population outperform those in the conditions with mere conflicts disclosures. Further research is necessary to identify contexts in which biases can be calculated reliably, and to understand how to best communicate that information to laypersons so that it is useful to them.

This Article also explored mechanisms for tailoring disclosure mandates to particular subpopulations that actually receive biased advice. Analysis revealed that a mandate to disclose conflicting interests can hurt the potentially large proportion of laypersons who are nonetheless receiving accurate advice. Thus, disclosure mandates should not be imposed unless there is particularized evidence of an actual advisor being biased, and then disclosure mandates should be tailored to the particular laypersons receiving biased advice. On the other hand, the present study demonstrated that even when an advisor has aligned interests, a disclosure helps laypersons place their reliance and improve performance. Affirmative disclosures can help with a trust deficit.

A primary finding of the present study is that a disclosure mandate improves layperson performance when unbiased advice is available too, as may be true in many market settings. A second opinion from an unbiased advisor is a much better remedy for biased advice than disclosure. Indeed, disclosure of conflict plus a second opinion from an unbiased advisor helps laypersons perform as well or better than simply providing accurate advice in the first place. Still, it bears emphasis that this is just a complicated way of rectifying the problem that the original advisor had conflicting interests.

Notwithstanding the love for market-based solutions amongst both scholars and politicians, this study strikes a pessimistic note, given its findings about laypersons' self-assessments. This study found an inverse relationship between laypersons' accuracy and their own confidence in their performance. The present study suggests that policymakers should give increasing attention to policy mechanisms that align the interests of advisors and laypersons, and that channel laypersons toward unbiased advice, which is the strongest determinant of layperson performance.

METHODOLOGICAL APPENDIX

This Appendix provides details about the methods used in the experiments. Human subjects were recruited from e-mail lists and websites nationwide, including Craigslist, Facebook, and Amazon Mechanical Turk, to complete the study hosted on a third-party website.¹⁶³ Participants completed an online informed consent form approved by Harvard University's Institutional Review Board. The 198 participants recruited from Mechanical Turk were paid \$0.10 to \$0.15 each to complete the study, in addition to an accuracy-based \$100 prize drawing. The remaining 545 participants received no payments for participation but were eligible for a \$100 prize for accuracy. All the subjects were told: "The person who gets closest to the actual value most often wins the \$100 prize. So try your best to be accurate."

As shown on Table 1, the present study has replicated the findings from CLM's classroom-based study, which thereby calibrates the present study's experimental design, making subsequent findings commensurate. Three experimental conditions were nearly identical to those tested in the CLM study, though the conditions were renamed for consistency with the other conditions tested here:

- 1UA (with one unbiased advisor and a disclosure that interests are aligned, corresponding to CLM's "accurate" condition);
- 1BN (with one biased advisor and no disclosure, corresponding to CLM's "high/undisclosed" condition); and
- 1BC (with one biased advisor and a disclosure of conflicting interests, corresponding to CLM's "high/disclosed" condition).¹⁶⁴

Although the standard deviations are higher in the present study, the point estimates for the means are quite similar across the CLM study and the present

¹⁶³ See Gabriele Paolacci et al., *Running Experiments on Amazon Mechanical Turk*, 5 JUDGMENT & DECISION MAKING 411 (2010) (describing the increasing use of Mechanical Turk by social scientists).

¹⁶⁴ See Cain et al., *supra* note 32, at 10. In addition to the methodological differences noted above (i.e., an online study versus a classroom study, and no participants assigned to the advisor role), there was one other difference between the CLM study and the present study. In CLM, the final three jars were "feedback rounds" in which the actual value of the jars was revealed to estimators after they rendered their estimates. CLM found no significant effects from this feature. *Id.* at 18. The feature was excluded here partly because of a concern that participants would communicate the right answers to future participants who might learn of the study through social networking sites.

study, with no more than a statistically insignificant difference of \$0.33 for comparable conditions.

This cross-study comparison should lend additional credence to both the CLM study and the present study, and has methodological interest, since it helps validate these two different approaches to behavioral research. Cain and colleagues paid a relatively homogenous group of Carnegie Mellon University students an average of \$10 each to participate,¹⁶⁵ while the present study recruited participants nationwide at an average cost of only \$0.18 each, including both the per-person payments (zero to \$0.16 each) and the \$100 prize drawing. The fact that this study has replicated a classroom study, and has done so with arguably broader external validity at one-fiftieth of the cost per participant, is promising for the future of empirical legal studies.

The participants in the present study were 72% white/Caucasian. Approximately a third had “some college” for their highest educational level, and another third had graduated from college. The mean age was thirty-three, with only about a quarter being the college age of eighteen to twenty-two years. Thus, this sample is somewhat more heterogeneous and more representative of the American population than the CLM sample, though it is still far from a demographically valid sample.

Table 2 shows the photographs of six jars used in this study (at reduced size), along with the actual values of the coins in those jars,¹⁶⁶ the mean personal estimates rendered by advisors in the accurate condition,¹⁶⁷ the mean advice given in each condition,¹⁶⁸ and the mean estimates rendered by unadvised laypersons in each condition (from the present study).

Table 3 summarizes the conditions employed in this study, manipulated according to the number of advisors (zero, one, or two), the quality of the advice (accurate, biased, or even more biased because of a disclosure mandate), and the type of disclosure given (none, disclosure of conflicting interests, or disclosure of average bias). Table 3 also lists the number of participants in each condition (n), the primary dependent variable used in the study, which is the mean inaccuracy defined in terms of virtual error,¹⁶⁹ and the standard deviation (SD).

¹⁶⁵ See *id.* at 9.

¹⁶⁶ See *id.* at 14 tbl.4.

¹⁶⁷ See *id.*

¹⁶⁸ See *id.* at 13, 15 tbl.5.

¹⁶⁹ See discussion *supra* note 70.

Table 4 reports the actual errors, in contrast to the “virtual errors” discussed in the body of this Article, along with laypersons’ self-reported confidence in their estimates, by condition.

Table 1:
Comparison of Layperson Virtual Error in CLM and Present Study

Condition	CLM Study			Present Study			Difference of Means (<i>p</i> value)
	<i>n</i>	mean	SD	<i>n</i>	mean	SD	
accurate 1UA	27	3.41	1.36	39	3.56	2.64	-0.15 (.78)
high disclosed 1BC	27	6.20	2.62	116	6.49	3.28	-0.29 (.67)
high undisclosed 1BN	26	4.52	1.58	43	4.85	2.65	-0.33 (.57)

Table 2:
The Experimental Stimuli

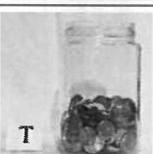
	Actual Value: \$10.01 Advisors' Personal Estimate: \$11.85 Accurate Advice: \$12.30 High-Disclosed Advice: \$17.20 High-Undisclosed Advice: \$16.20 Unadvised Laypersons' Estimate: \$3.22
	Actual Value: \$19.83 Advisors' Personal Estimate: \$16.73 Accurate Advice: \$16.80 High-Disclosed Advice: \$22.25 High-Undisclosed Advice: \$18.90 Unadvised Laypersons' Estimate: \$7.55
	Actual Value: \$15.58 Advisors' Personal Estimate: \$12.75 Accurate Advice: \$14.00 High-Disclosed Advice: \$25.25 High-Undisclosed Advice: \$15.75 Unadvised Laypersons' Estimate: \$6.95
	Actual Value: \$27.06 Advisors' Personal Estimate: \$18.39 Accurate Advice: \$20.00 High-Disclosed Advice: \$27.75 High-Undisclosed Advice: \$24.90 Unadvised Laypersons' Estimate: \$10.45
	Actual Value: \$24.00 Advisors' Personal Estimate: \$21.30 Accurate Advice: \$21.50 High-Disclosed Advice: \$28.25 High-Undisclosed Advice: \$25.30 Unadvised Laypersons' Estimate: \$9.64
	Actual Value: \$12.15 Advisors' Personal Estimate: \$13.07 Accurate Advice: \$14.25 High-Disclosed Advice: \$24.25 High-Undisclosed Advice: \$19.90 Unadvised Laypersons' Estimate: \$5.22
<u>Average Across All Jars</u>	Actual Value: \$18.16 Advisors' Personal Estimate: \$15.68 Accurate Advice: \$16.48 High-Disclosed Advice: \$24.16 High-Undisclosed Advice: \$20.16 Unadvised Laypersons' Estimate: \$7.19

Table 3: Summary of Conditions and Results for Layperson Virtual Inaccuracy

Purpose of Experimental Condition	Label	# of Advisors & Advice Quality	Disclosure Type	Location of Disclosure	n	Mean Virtual Inaccuracy (SD)
calibrate online study with CLM; establish benchmark for layperson performance	1UA	1 unbiased	aligned interests	after advice	39	3.56 (2.64)
calibrate online study with CLM; test impact of bias on layperson performance compared to 1UA	1BN	1 biased	nothing	N/A	43	4.85 (2.65)
calibrate online study with CLM; test impact of disclosure mandate compared to 1BN	1BC	1 very biased*	conflicted interests	after advice	116	6.49 (3.28)
measure episodic asymmetry, test potential impact of a policy that would ban conflicted advice	No Advisors	0 N/A	N/A	N/A	42	9.76 (2.88)
test sequencing of disclosure vs. 1BC to test anchoring effects	1BCF	1 very biased*	conflicted interests	first (before advice)	114	6.25 (3.18)
test disclosures of actual bias (along with disclosure of conflicts as in 1BC) as potential policy improvement	1BCB	1 very biased*	conflicted interests & average bias	after advice	56	6.95 (4.79)
test sequencing of disclosure of actual bias	1BCBF	1 very biased*	conflicted interests & average bias	first (before advice)	66	6.74 (4.50)
test disclosure mandate on laypersons receiving accurate advice; explore tailored disclosures policy	1UC	1 unbiased	conflicted interests	after advice	53	4.77 (3.03)
test effect of aligned-interest disclosure used in 1UA as potential mechanism for improving reliance	1UN	1 unbiased	Nothing	N/A	52	4.71 (2.75)
test second opinion as remedy for biased advice	2N	1 biased, 1 unbiased	nothing, nothing	N/A	57	3.97 (3.15)
test disclosure mandate in multiple advisors situation, compared to 2N	2D	1 very biased*, 1 unbiased	conflicted interests, aligned interests	after advice	53	2.75 (1.96)
test sequencing of advisors vs. 2D	2DR	1 unbiased, 1 very biased*	aligned interests, conflicted interests	after advice	52	3.56 (2.12)

* Advice given was equal to that given in CLM's "high-disclosed" condition; other biased advice was equivalent to CLM's "high-undisclosed" condition.

Table 4:
Actual Errors and Layperson Confidence

Condition	<i>n</i>	Layperson Actual Error (estimate minus truth)		Layperson Confidence	
		Mean	SD	Mean	SD
1UA	39	4.88	2.72	4.77	1.90
1BC	116	6.42	3.51	5.10	1.89
1BN	43	4.95	3.47	5.38	1.82
1UC	53	6.35	3.42	5.25	1.89
1UN	52	6.12	3.42	5.40	1.61
1BCB	56	6.67	4.00	5.18	2.17
NoAdvisors	42	11.65	3.84	6.07	2.01
2N	57	4.82	3.15	4.79	1.98
2D	53	3.47	1.24	4.69	2.21
2DR	52	4.24	2.46	5.14	1.96
1BCF	114	5.95	3.18	5.13	1.93
1BCBF	66	6.29	3.42	4.87	2.23

87

Who's the Fairest of Them All? A Comparative Analysis of Financial Advisor Compensation Models

by John H. Robinson

John H. Robinson is branch manager and managing director of a Honolulu, Hawaii-based wealth management practice. He has been in the industry since 1989.

"Incentives are the cornerstone of modern life. And understanding them—or, often, ferreting them out—is the key to solving just about any riddle, from violent crime to sports cheating to online dating."

—Steven D. Leavitt and Stephen J. Dubner,
Freakonomics

Retail investors obtain personalized financial guidance from a variety of sources, including national and regional brokerage firms, registered investment advisory firms (RIAs), and independent financial planners. Depending on the source, there are essentially three ways investors pay for financial guidance: commissions, asset-based fees, and flat fees (such as retainers and hourly fees). Over the past decade or so, as the investment industry has evolved from its sales-oriented origins toward a financial planning orientation, the traditional commission model has come under increasing fire for the seemingly obvious conflict of interest that commission-based sales create between the financial advisor¹ and the investor. As a result of this criticism, there has been a major shift away from transactional compensation toward asset-based fees, and, to a lesser extent, toward flat-fee planning—the rationale being that fee-based compensation better aligns advisor and investor interests. In fact, some advisors have adopted an almost moralistic position in

Executive Summary

- There are three primary modes by which investors pay for financial advice: commissions, asset-based fees, and flat fees. This paper examines the economic incentives at work in each and suggests that the debate over which model is "fairest" is flawed because all three models contain incentives that can lead to conflicts of interest and each may also represent an optimal choice for certain investor circumstances.
- The recent trend away from commissions in favor of fee-only planning may not represent a best-practice model for the profession. Alternatively, the ideal compensation platform may be one that incorporates all three models.
- While conflicts of interest in the commission model are seemingly obvious, data exist to suggest that the impact of these conflicts may be overstated and that the commission model may have a cost advantage for some investors.
- A clear advantage of asset-based fees is

that advisor compensation is tied to performance. Still, conflicts in this model may arise from inherent disincentives to recommend strategies that lead clients to reduce assets under management, even if such strategies are in the clients' best interests.

- The flat-fee model is the only one that truly allows the client to pay for broad-based financial planning guidance that is not merely incidental to the investment plan. Nonetheless, shirking and over-billing are potential conflicts of interest that arise under this model.
- There is little evidence to suggest that regulatory differences lead fee-based advisors to be either more qualified or to act more ethically than commission-based advisors; however, the fee-based models are clearly superior with respect to fiduciary disclosure requirements. It can be argued that regulatory inequality denigrates the commission model's credibility.

advocating the exclusive use of either of the fee-based models. This position is regularly echoed in both the mainstream media and in academic financial journals. So vocal is the fee-only movement that there has been very little formal discussion of limitations associated with either of the fee-based models or comparative analyses

of the merits or disadvantages of all three models relative to each other. But are the two fee-based models always better for the investor? This paper addresses this question by ferreting out and examining some of the underlying economic incentives at work in all three models and by identifying those that may

create advisor/investor conflicts. The analysis concludes that all three models contain incentives that align advisor and client objectives, and which can create significant conflicts of interest. This somewhat controversial finding has at least three important implications.

1. No single compensation model ensures investors that the advice they receive will be entirely unbiased and objective.
2. Since all three models have limitations, the financial planning community should think carefully about promoting fee-only planning as a best-practices model for the profession.
3. Since each model may also be an optimal choice for certain sets of investor circumstances, the "fairest" compensation platform may be one that offers investors the ability to choose from all three models.

The Commission Model

Despite the widespread migration toward fee-based compensation models, thousands of Series 7 registered financial advisors at hundreds of broker/dealers nationwide still provide commission-based guidance.² Commission-based advisors are regulated by the Securities and Exchange Commission (SEC) under the Securities Exchange Act of 1934 and by various self-regulatory agencies, most notably the National Association of Securities Dealers (NASD).

As noted above, the primary criticism of the commission model stems from the obvious problem that the advisor's economic incentives appear to be, at best, detached from or, at worst, in direct conflict with the interests of the investor. This conflict arises from the simple fact that the advisor receives a commission regardless of whether the investments recommended succeed or fail and because the commission model creates an economic incentive for the advisor to steer investors toward products that pay the highest commissions, even if those products are not necessarily

best suited for the investor. Further, advisors may have an incentive to generate transactions purely for the sake of generating income for themselves rather than for the benefit of the client. As such, detractors of the commission model suggest that the quality of financial guidance dispensed from commission-based advisors should be viewed with skepticism and should be discounted relative to that of fee-based advisors.

As straightforward and as convincing as these criticisms of the commission model may seem, an objective comparison of the three models requires us to look beyond the obvious and to consider whether there are any underlying incentives or circumstances that might actually make the commission-based model a better choice for certain investors than the fee-based models. Specifically, we must ask whether the conflict-of-interest argument accurately reflects the behavior of financial advisors.

Commission-based advisors contend that, in reality, their interests are far more aligned with investors than one might think. From their perspective, the financial services industry is a crowded and competitive place, and attracting new clients is more difficult than ever. If one accepts the assumption that investors will eventually leave advisors who abuse their trust, then commission-based advisors may indeed have an incentive to forgo short-term gratification from commission maximization in order to keep investors happy over the long run so that they will continue to generate revenue and refer other investors.

Similarly, commission-based advisors also argue that they do not have an incentive to steer clients disproportionately toward investments with high upfront commissions because such investments tend to be "one-shot deals" (that is, the

advisor may not get paid again). Allocating client assets entirely to these products would mean that the advisor must constantly be prospecting for new investors.

Although such an economic defense is likely to be met with skepticism from the fee-only camp, there is evidence to support

“So vocal is the fee-only movement that there has been very little formal discussion of limitations associated with either of the fee-based models or comparative analyses of the merits or disadvantages of all three models relative to each other.”

the commission-based advisor's position, and to suggest that the conflict-of-interest criticism surrounding the commission model may indeed be overly simplistic. If commission-based advisors truly have a greater economic incentive to put their interests ahead of investors, then one would expect commission-based advisors to have higher returns on assets (ROA) than their fee-based counterparts. In fact, the opposite is true—and by a rather wide margin. Industry data regularly report that the average ROA for brokers is below .75 percent, while the average ROA for independent RIAs is approximately 1.3 percent and rising.³

Not only does this suggest that commission-based advisors as a group are not solely motivated to maximize commissions, it implies that the commission-based model may have a significant cost advantage over fee-based investing. Further, there is some evidence that contradicts the assertion that commission-based advisors tend to steer clients toward products with higher commissions. One example is the recent trend among brokerage advisors away from commission-based products toward non-advisory fee-based accounts.

While this trend has been at the center of a number of controversies, one major regulatory concern surrounding these accounts has been the issue of “reverse churning”—the notion being that commission-based advisors are shifting relatively inactive accounts into fee-based accounts in order to increase revenue. Thus, it appears that at least some commission-based advisors view fee-based planning as more lucrative than commission-based planning (that is, advisors may have an economic incentive, rather than solely a moral incentive, to migrate toward fee-based platforms). From this, one might also conclude that there are certain investor sets that would be better served—at least from a purely economic perspective—by the commission model. Intuitively, buy-and-hold stock portfolios, breakpoint-eligible mutual fund portfolios, static or laddered bond or certificate of deposit portfolios, and accounts with sizable cash balances are all examples of portfolios that might be better served under a commission-based model than a fee-based model. Similarly, assuming that the advisor has the proper qualifications, experience, and designations, a commission-based advisor might be a comparatively inexpensive resource for incidental, modular financial planning guidance.

The Asset-Based Fee Model

Financial advisors who are compensated for financial planning and investment advice via asset-based fees are registered as investment adviser representatives (IARs) under the Investment Advisers Act of 1940, and are required to hold the NASD’s Series 65 registration. Depending on the total amount of client assets under management, registered investment advisory firms (RIAs) must be registered either with the SEC (RIAs with greater than \$25 million in client assets) or their state’s securities regulator. Although most national and regional investment firms carry dual broker/dealer and RIA registrations, the strongest growth in the RIA arena in recent years has come from fee-only RIAs,

which operate independently from any broker/dealer. Of the more than 10,000 SEC-registered RIAs (a group that includes hedge funds, institutional and global money managers, pension consultants, and mutual portfolio managers), it is estimated that more than 3,000 are small, independent firms engaged in providing financial planning services.⁴ The majority of the more than 13,000 state-registered RIAs fall into this category as well.⁵

The most vociferously touted benefit of the asset-based fee model is that it aligns the advisor’s economic incentives with those of the investor. This seems intuitive because, under the asset-based fee model, the advisor is paid more if the portfolio value rises and less if the value falls. But are there disadvantages or conflicts of interests with the asset-based fee model, too? We’ve already discussed as one potential conflict the notion that some advisors might adopt this model in order to earn higher revenue streams for themselves than they would under the commission model. But are there others?

Again, we begin by examining the model’s incentive structure. At first blush, it seems intuitive that advisor and client interests are always aligned under the asset-based fee model. But there are at least two scenarios in which advisor guidance under this model may be directly at odds with the interests of the client.

First, since the asset-based fee advisor is paid based on a percentage of assets under management, it can be argued that there is a strong disincentive to provide investment solutions that do not involve advisor management or that might reduce the amount of investor assets under management. For example, the asset-based fee advisor might be disinclined to advise investors to reduce debt or invest in assets such as real estate or art.⁶ Similarly, an asset-based fee advisor might be reluctant to recommend holding

liquid assets that require little ongoing management, such as cash or static/laddered bond portfolios outside of the fee arrangement, even if it is in the investor’s best interests to do so.

Second, since most asset-based fee advisors eschew commissions, it is reasonable to expect there to be a strong disinclination to recommend certain products such as life insurance, long-term care insurance,

“At first blush, it seems intuitive that advisor and client interests are always aligned under the asset-based fee model. But there are at least two scenarios in which advisor guidance under this model may be directly at odds with the interests of the client.”

disability insurance, or certain annuity contracts that are only available on a commission basis, even if these products may be appropriate for the investor’s circumstances. While it is unlikely that these scenarios would be considered a breach of fiduciary duty, the economic incentives in these examples indisputably place asset-based fee advisors’ interests at odds with those of their clients.

From the investor’s perspective, the notion that both the advisor and client benefit when asset values rise, and suffer when assets fall, is indeed a desirable attribute and is a distinct advantage of the asset-based fee model for certain investor circumstances. Intuitively, the model seems particularly well suited for long-term, growth-oriented investors who prefer a hands-off approach to investing (that is, advisor discretion) and for actively traded accounts. Managed accounts and wrap-fee no-load or load-waived mutual fund portfolios also seem to lend themselves well to this model.

As with the commission model, asset-based-fee advisors who incorporate financial planning guidance in their service offering may also add value to the relationship. But if these advisors promote fee-only planning, it follows that they should direct clients to other professionals to obtain certain commission-based insurance products that may be needed to address common life event risks, such as disability insurance, survivorship life insurance, and long-term care insurance.

The Flat-Fee Model

Incentives in the flat-rate model are difficult to analyze because fees are levied in multiple formats. Two of the most common formats are retainers and hourly billing. Retainers typically are billed quarterly and are often based on the client's net worth rather than portfolio value. Hourly fees tend to be levied by advisors hired to prepare a one-time plan or who are employed by clients who wish to concentrate on one or two specific issues (such as college funding, complex tax planning, or stock option analysis) rather than paying a steeper fee for comprehensive financial planning services.

Under the commission and asset-based fee models, the advisor is compensated specifically for investment advice, while all other planning guidance is effectively incidental to the relationship. In contrast, an undeniable benefit of the flat-fee model is that it truly allows the advisor to be compensated for financial planning advice. As such, flat-fee advisors argue that they can be entirely objective in their recommendations. For example, if a flat-fee advisor believes a client's portfolio is overweighted in securities and underweighted in real estate, he or she is free from conflicting economic incentives in recommending the allocation change. It is for this reason that some advisors regard the flat-fee model as ethically superior to the other compensation models.⁷ But is this model, as its proponents claim, truly free from conflicts of interest and is it truly the best choice for all investors?

Quantitative statistics, such as returns-on-assets data, are more ephemeral under the flat-fee practices because of the varied methods by which fee-based advisors charge for their services and the fact that flat-fee advisors still make up a relatively small portion of the overall marketplace. But some insight into the underlying economic incentives in the flat-fee model may be gleaned by analogy from examining issues of ethical contention that arise with compensation practices in the legal profession. In the practice of law, retainers and hourly billing are the most common billing methods among defense attorneys and transactional lawyers (contingency fees are more common among plaintiff attorneys). In the legal profession, a common criticism of retainers is that it creates an economic incentive for shirking (that is, to do as little work as possible to justify the fee). For example, a tobacco company may hire a prominent law firm on retainer to defend against individual plaintiff lawsuits. To maximize profitability, the law firm may be inclined to assign the work to lower paid, less experienced associates or to use templates from previous cases. The same incentives likely apply to retainers in flat-fee advisory practices. Because asymmetries of knowledge exist, flat-fee advisors may have an incentive to overstate the amount of work they have done for clients. Similarly, a controversial issue in the legal profession regarding hourly fees is the concept of "value billing." An attorney may have spent dozens of hours researching a particular issue. But once this research has been done once, it may be applied to the same type of case over and over again. Should the attorney bill the same number of hours for each subsequent client, or should he or she bill the actual time spent? By analogy, it may take an hourly planner dozens of hours to prepare a customized, comprehensive financial plan for a client. But once the basic tem-

plate for the plan is in place, it may be possible to copy the format along with certain boilerplate explanations of concepts from one plan to the next. Thus, there may be an incentive to "pad" bills as an hourly practice becomes more systematized.

In addition, just as defense attorneys are paid regardless of whether they win or lose, a potential disadvantage of the flat-fee model is that advisors may be less accountable for their guidance, at least as it applies to the securities portion of the portfolio. Whereas both commission-based and asset-based advisors have a clear economic incentive to closely monitor investment portfolios, the incentive is less direct for flat-fee advisors. In this regard, the conflict-of-interest issue that arises with flat-fee advisors may be similar to the reverse churning issue with commission-based advisors who encourage clients to adopt non-advisory, asset-fee-based brokerage accounts. If a flat-fee advisor recommends few investment changes from one year to the next, the question may arise as to whether the lack of change is due to the belief that the portfolio remains sound or due to shirking on the part of the advisor.

Another potential disadvantage of the flat-fee model is cost, insofar as clients may be forced to pay more to implement a plan than they would under the other two models. For example, if the flat-fee advisor

“An undeniable benefit of the flat-fee model is that it truly allows the advisor to be compensated for financial planning advice.”

recommends certain products such as certificates of deposit, life insurance, or annuities, the client may be subject to the same commissions, expenses, or surrender periods as if they had not paid the advisor a fee. Conversely, as with fee-only asset-based advisors, flat-fee advisors may be

inclined to avoid recommending certain commission-based products, even if these products might be the best tools to address a client's needs or risk exposure.

To summarize, the flat-fee model offers a clear advantage over the other two models by allowing advisors to be paid for comprehensive financial planning guidance instead of just for investment guidance. In doing so, it can be argued that advice under the flat-fee model may be more objective because advisors do not have an economic incentive to encourage clients to keep assets in securities portfolios. But the corollary of this thesis is that flat-fee advisors may be less accountable for investment performance and may have a greater incentive to shirk than advisors under the other two models. In addition, both flat-fee and fee-only asset-based advisors face challenging decisions with respect to recommending certain products that are only available on a commission basis.

Based on these findings, one could conclude that flat-fee planning is particularly well suited for people for whom objective comprehensive financial planning guidance is valued more highly than traditional investment planning. It may also be well suited for people who are seeking a one-time financial planning analysis or a second opinion on an existing plan. Hourly planning also seems to be a logical solution for people seeking guidance on one or two specific complex planning issues such as tax planning, college funding, executive compensation issues, or stock option analysis.

The Impact of Regulation on Advisor Incentives

Ostensibly, one of the primary purposes of regulation in the investment industry is to build public trust by using rulemaking and enforcement to create incentives for advisors to do what is right for their clients. Thus, a comparison of advisor compensation models would be incomplete without considering the impact of regulation on incentives under each of the three models. Regulatory differences stem from the fact that Series 7-registered, commission-based

advisors are governed by the Securities Exchange Act of 1934, while Series 65 IARs (both asset fee and flat fee) are governed by the Investment Advisers Act of 1940. Under the Securities Exchange Act, commission-based advisors are held only to a standard of suitability in their recommendations, while fee-based advisors are held to fiduciary standards that (1) require them to place the interests of the client first and (2) hold the advisor personally liable for breaches of fiduciary duty. The notion that fiduciary responsibility is a higher ethical standard than mere suitability has been central to the position of many fee-only proponents. But commission-

“Because asymmetries of knowledge exist, flat-fee advisors may have an incentive to overstate the amount of work they have done for clients.”

based advisors counter that regulation has evolved considerably since 1934, and that there is no longer any practical difference between the two standards because the rules imposed by the NASD have become so stringent that they force advisors to act with integrity. Further, commission-based advisors argue that client assets may actually be safer in the broker/dealer environment than with an RIA because of SIPC (Securities Investor Protection Corporation) coverage and because of steep firm liability in the instance of fraud.

Thus, to continue our analysis requires investigating whether the client suitability/fiduciary difference truly creates an incentive for fee-only advisors to act more ethically than their commission-based counterparts, or whether the value of this difference has been exaggerated.

On the surface, the threat of personal liability under the 1940 act does not appear to be a significant deterrent to unethical behavior. There certainly are scores of

examples of rogue brokers who have churned clients' accounts, engaged in unauthorized trading, failed to properly disclose risks and expenses, embezzled, and otherwise violated their clients' trust. On the other hand, there are abundant examples of fee-only advisors who have breached their fiduciary duty by receiving undisclosed commissions or kickbacks from sales of products such as index annuities, private equity schemes, limited partnerships, and viatical settlements, or who have defrauded their clients through complex Ponzi schemes. For every Frank Crutadauria there appears to be a Bradford Bleidt.⁹

Contrary to popular perception, the Series 65 designation does not appear, in and of itself, to automatically confer a higher degree of ethical behavior on IARs. In fact, there is strong evidence to suggest that the deterrent incentive of the fiduciary liability

threat is weak due to inadequate enforcement resources. In the mid-1990s, then SEC chairman Arthur Leavitt famously quipped that, based on 81 SEC staffers and 22,500 RIAs, an advisor could expect to be audited once every 44 years. To address this oversight problem, Congress passed the National Securities Improvement Act of 1996. This law shifted oversight of all RIAs with under \$25 million in assets to the individual state securities commissioners. Although this law was embraced by the states and has been successful to a degree, today there are more than 10,000 SEC RIAs and more than 13,000 state regulated RIAs. In 2005, the SEC's Office of Compliance Inspections and Examinations (OCIE) conducted 1,530 examinations, of which 932 were "routine" audits.⁹ Based on these figures, an SEC RIA might expect to be audited once out of every ten years. Evidence of inadequate enforcement resources at the state level can be found in the comments

of Consumer Federation of America director Barbara Roper who noted that "states are doing the best they can with grossly inadequate resources. In too many cases, however, their best still leaves investors dangerously vulnerable to fraud."¹⁰

Another important difference between the 1934 Securities Exchange Act and the 1940 Investment Advisers Act pertains to disclosure. Under the 1940 act, IARs are required to disclose all fees they will receive as compensation for their services. In contrast, disclosure rules under the 1934 act are considerably more abstruse. While commissions on stock trades are generally disclosed and sales charges on mutual funds must be disclosed via prospectus, advisor compensation on products such as bonds, CDs, annuities, and various insurance lines are generally opaque to investors.

While the aforementioned ROA statistics do not suggest that lack of disclosure has

led to widespread abuse, the fact remains that it is possible for advisors to choose to recommend one product over another based on commission.

Evidence that some advisors select products based on the commissions they pay is found in the way in which many products are marketed to them. For example, a full-page ad for a particular annuity contract in the November 2005 issue of the *Journal of Financial Planning* touts a 10 percent commission as a primary benefit of the product. Although some commission-based advisors may be inclined to point out the recent controversies surrounding RIAs and the conflicts of interest arising from undisclosed hard- and soft-dollar fee

arrangements, on balance, disclosure under the fee-based models seems more transparent than under the commission-

“A potential disadvantage of the flat-fee model is that advisors may be less accountable for their guidance, at least as it applies to the securities portion of the portfolio.”

based model. In terms of economic incentives, it can be reasonably inferred that if commission-based advisors were required to disclose the amount they were paid from the sale of each product, commissions on certain products would fall and commissions would generally become

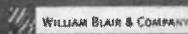
Thank you
to our
sponsors

Presenting:

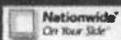


Muller-Kump & Company, Inc.
Investment Management

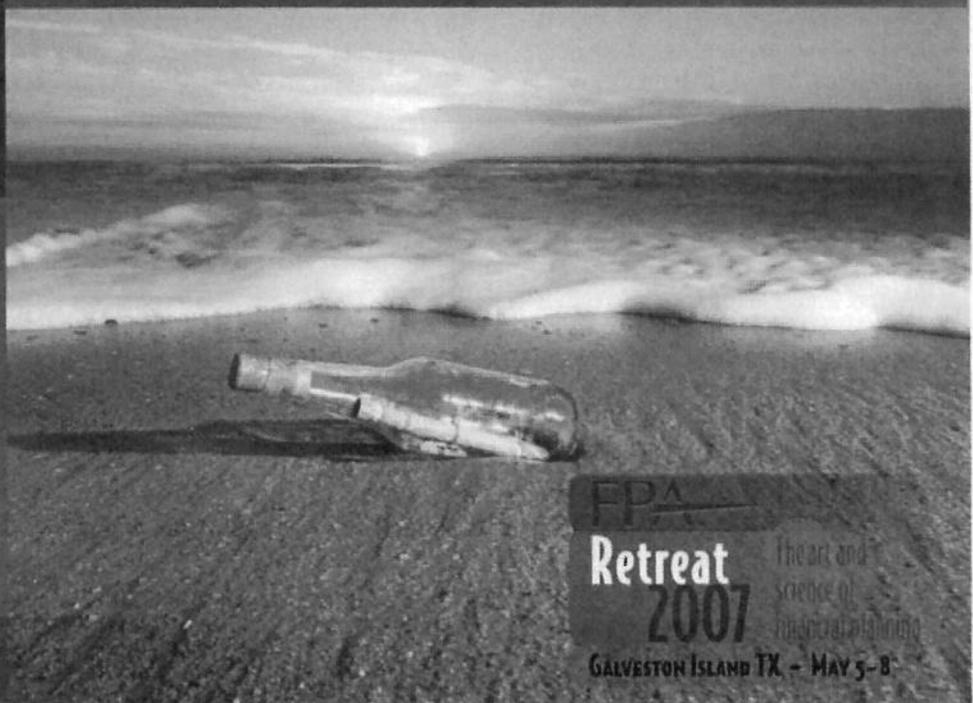
Associate:



National:



for another great event!



FPA
Retreat
2007

The art and
science of
financial planning

GALVESTON ISLAND TX • MAY 5-8

more level across product lines. In this regard, regulatory incentives seem to favor the two fee-based models.

“There is a widely held perception that fee-based advisors are more qualified to provide financial planning guidance than commission-based advisors. But nothing in the 1940 act or the licensing criteria suggests that Series 65 holders are appreciably more qualified than Series 7 holders.”

Finally, there is a widely held perception that fee-based advisors are more qualified to provide financial planning guidance than commission-based advisors. But nothing in the 1940 act or the licensing criteria suggests that Series 65 holders are appreciably more qualified than Series 7 holders. In fact, according to a 2005 survey conducted by the College for Financial Planning, 56 percent of CFP certificants report receiving income from a combination of fees and commissions. Furthermore, according to the CFP Board, of the approximately 49,000 CFP certificants, 37 percent (18,500) are affiliated with the top 30 broker/dealers, insurance companies, and mutual fund companies. Thus, educational qualifications do not appear to weigh in as an advantage of one model over another.

Analysis Implications

To borrow a quotation from the best-selling book, *Freakonomics*, by Steven D. Levitt and Stephen J. Dubner, “Morality, it could be argued, represents the way that people would like the world to work—whereas economics represents how it actually does work.” With that in mind, this paper has sought to examine the economic incentives at work in the three major advisor compensation models in order to attempt to deter-

mine if one model is clearly a superior choice for investors. In doing so, we have found that economic incentives exist in all three models that can lead to conflicts between advisors’ interests and those of their clients, and that each model has certain unique attributes that may make it the “fairest” choice for certain sets of investors.

The major implication of this finding is that, contrary to public opinion, the promotion of fee-only planning may not be a best practice for the industry. Alternatively, it is logical to conclude that a best-practice standard might be a platform that offers clients a choice of some combination of all three models. A further implication of this work is that professionals who wish to adopt the three-model platform must be dual-registered with both a broker/dealer and an RIA.

This comparative analysis also included an examination of incentives imposed by regulation under the three models. In doing so, surprisingly little evidence was found to suggest that fiduciary liability is effective in deterring fraud and creating higher ethical conduct among fee-based advisors. But a review of the literature reveals that there is indeed a widely held perception that fiduciary responsibility is a higher ethical standard than the suitability standard for commission-based advisors. This suggests that the credibility of commission-based advisors is being artificially denigrated by unbalanced regulation.

This perceived imbalance could be alleviated by adopting legislation that would apply the fiduciary standard to broker/dealers as well as to RIAs. To date, however, the brokerage industry has lobbied against

this change, presumably out of fear that it would lead to increased liability. Interestingly, many individual commission-based advisors favor such a regulatory change because they believe it would help to level the competitive playing field. In a recent survey conducted by the Financial Planning Association, more than 86 percent of respondents who were wirehouse advisors indicated they would support the adoption of a fiduciary standard. More telling was the fact that there was little difference in support for the change between wirehouse advisors who hold the CFP certification and those who do not.¹¹

In terms of other major regulatory incentives, this paper finds that the more stringent compensation disclosure requirements under the Investment Advisers Act of 1940 favor the two fee-based models, and that a lack of transparency under the Securities Exchange Act of 1934 still fosters an environment where commission-based advisors have an incentive to choose products for their clients based on the commissions they receive rather than on the merits of the product for the client. It seems intuitively predictable that if commission-based advisors were required to disclose the amount of commission they receive on each transaction, commissions would fall and commissions would become more level across product lines. As long as such informational asymmetries regarding compensation are permitted to exist between advisors and their clients, the commission-based model may be viewed as inferior to the two fee-based models. Thus, while each of the three compensation models has its unique advantages and disadvantages, improved transparency and adoption of the fiduciary standard would go a long way toward improving the image of the commission-based model relative to asset-based and flat-fee models.

In summary, although the conclusions drawn from this comparative analysis of the incentives in each of the three major advisor compensation models may be viewed as controversial, particularly to staunch fee-only and flat-fee advocates, a

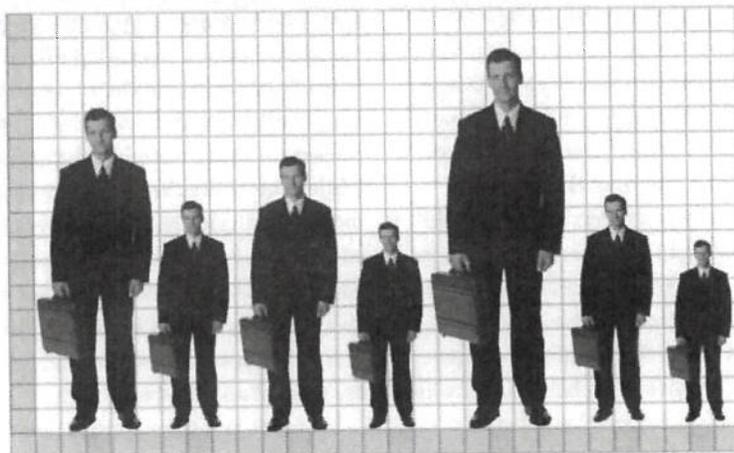
primary purpose of this exercise has been to spur healthy debate on the subject. It is hoped that this paper will serve as a springboard for further discussion.



Endnotes

1. While it is understood that the term "adviser" refers specifically to investment adviser representatives (IARs) and the term "advisor" is used to refer broadly to all those offering investment guidance (that is, investment advisers and registered representatives collectively), due to the frequent references to both throughout this paper, "advisor" is generally used.
2. Although broker/dealers are permitted to offer fee-based compensation programs that are exempt from the Investment Advisers Act of 1940, the debate over the legitimacy of this practice is regarded as a separate issue and, as such, discussion of this topic has intentionally been omitted.
3. "Evolution Revolution 2005—A Profile of the Investment Adviser Profession," a report prepared by a joint collaboration of the Investment Adviser Association and National Regulatory Services.
4. "Evolution Revolution 2006—A Profile of the Investment Adviser Profession," a report prepared by a joint collaboration of the Investment Adviser Association and National Regulatory Services.
5. Ibid.
6. The same point could be raised with the commission model, though perhaps less so, since the commission advisor might benefit from a sales transaction, and he or she does not benefit from simply holding assets under management.
7. Nancy Opiela, "The Future of Fees," *Journal of Financial Planning*, August 2006.
8. Frank Gruttadauria was a Cleveland-based broker for Lehman Brothers who embezzled millions of dollars from clients over a 15-year period. Bradford Bleidt was a Boston based CFP certificant and RIA who defrauded millions of dollars from clients through a complex Ponzi scheme.
9. "Evolution Revolution 2006—A Profile of the Investment Adviser Profession," a report prepared by a joint collaboration of the Investment Adviser Association and National Regulatory Services.
10. Barbara Mallon, "Impact of the 1996 Reform Act on Investment Advisers," *The LawHost Online Law Journal*, 1998.
11. Duane Thompson, "Wirehouse Planners: FPA's Best-Kept Secret," *Journal of Financial Planning*, July 2006.

NEW FROM FPA THE POWERFUL BENCHMARK STUDY THAT WORKS FOR YOU



FPA.
FINANCIAL PLANNING ASSOCIATION
The Heart of Financial Planning™

McLAGAN
PARTNERS

When it comes to your financial planning practice, you need to manage, measure and benchmark.

Introducing the new, *FPA Practice Management Scorecard* provided by McLagan Partners funded in part by Fidelity Investments; a powerful and comprehensive tool you can use to benchmark the performance and growth of your individual practice in your backyard, where you live.

- Quick data template takes less than an hour to complete
- Totally customized for your practice
- Usable, actionable results
- Localized, relevant to your market

FPA members receive a discount of more than 20%, paying only \$195 per scorecard. Nonmembers pay \$245.

Register now! Deadline to participate is May 31.

Visit www.FPAnet.org/scorecard to begin the simple registration process. Results will be available August 20.

www.FPAnet.org/scorecard

88



Reputation and Product Quality

Author(s): William P. Rogerson

Source: *The Bell Journal of Economics*, Vol. 14, No. 2 (Autumn, 1983), pp. 508-516

Published by: RAND Corporation

Stable URL: <http://www.jstor.org/stable/3003651>

Accessed: 02-07-2018 18:16 UTC

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://about.jstor.org/terms>



JSTOR

RAND Corporation is collaborating with JSTOR to digitize, preserve and extend access to *The Bell Journal of Economics*

Reputation and product quality

William P. Rogerson*

This article considers the role that reputation plays in assuring product quality in markets where consumers can only imperfectly judge product quality even after consumption. Three conclusions are derived. First, high quality firms have more customers because they have fewer dissatisfied customers who leave and word-of-mouth advertising results in more arrivals. Second, higher fixed costs can result in a higher equilibrium level of quality. Third, the particular form that word-of-mouth advertising takes can have significant effects on the market outcome. Recommendations consisting of a report of whether the consumer intends to patronize the same firm again generate an externality that is absent when actual estimates of quality are communicated.

1. Introduction

■ This article models behavior in a market where product quality can vary and consumers cannot determine this quality before purchase. A growing literature has examined this type of market by employing the assumptions that a consumer can determine product quality immediately after purchase and that all other consumers become aware of his evaluation immediately or with a lag (Klein and Leffler, 1981; Shapiro, 1981; Dybvig and Spatt, 1982; von Weizsäcker, 1980). The consequence of these two assumptions is that there can be no misrepresentation of quality by firms over the long run. Consumers are, however, often capable of performing only very partial and vague evaluations of the quality of professional services they receive from doctors, lawyers, banks, mechanics, opticians, etc. Furthermore, the quality of a service from a given professional may vary from time to time. This combination of observer error and actual quality variance makes it difficult for consumers to evaluate correctly the quality of service that a firm produces. Furthermore, consumers cannot costlessly share all available information. In such a market a producer can contemplate staying in the market over the long run, even if he misrepresents the quality of his product.

This article derives three conclusions concerning the provision of quality in such a market. First, owing to the effect of reputation, high quality firms have more customers, because they have fewer dissatisfied customers who leave, and word-of-mouth advertising results in more arrivals. Second, higher fixed costs can result in a higher equilibrium level of quality. Thus, entry restrictions taking the form of mandatory fixed costs may well increase product quality in and of themselves. Third, the particular form that word-of-mouth advertising takes can have significant effects on the market outcome. In particular, recommendations consisting of a report as to whether the consumer intends to patronize

* Stanford University.

This work was supported by National Science Foundation Grant SES80-06654 at the Institute for Mathematical Studies in the Social Sciences, Stanford University. I would like to thank Robert Bates, Kathleen Hagerty, Alvin Klevorick, Roger Noll, Carl Shapiro, Alan Schwartz, Louis Wilde, and two anonymous referees for helpful comments and discussions.

the same firm again (i.e., recommendations are positive or negative) generate an externality that is absent in cases where actual estimates of quality are communicated.

2. Relative firm size

■ The purpose of this section is to argue that in equilibrium high quality firms will be larger than low quality firms.¹ Firms are assumed to be able to produce either high quality or low quality goods. Variables associated with high quality firms will be subscripted with an h ; those associated with low quality firms will be subscripted with an l . There are N identical consumers, each consuming one unit of the good per period. Consumers can only identify the quality of the good they receive after purchasing it, and even then only probabilistically. After purchasing a high quality good, consumers mistakenly identify it as low quality with probability α ; after purchasing a low quality good, consumers mistakenly identify it as high quality with probability β . Assume that $\alpha + \beta < 1$. This means that the consumer is more likely to believe that a good is of high quality if it is of high quality than if it is of low quality.

□ **The departure rule.** The consumer chooses a vector of real numbers $r = (r_1, r_2, r_3, \dots)$ and leaves his current firm at time t if and only if the fraction of high quality observations up to that point falls below r_t . Let B_h^t be the probability that a consumer spends at least t periods with a high quality firm. Define d_h by

$$d_h = \frac{1}{1 + \sum_{t=1}^{\infty} B_h^t}. \quad (1)$$

Let B_l^t and d_l denote the similar values for low quality firms. Proposition 1 summarizes the results on a firm's average number of departures if all its consumers use the same departure rule.

Proposition 1: Suppose that n new customers arrive at a firm in each period.

- (i) In the long run, the expected number of consumers at a high or low quality firm is, respectively, n/d_h or n/d_l .
- (ii) Therefore, in the long run a high or low quality firm experiences an average loss of, respectively, the fraction d_h or d_l of its customers per period.
- (iii) $d_h \leq d_l$.
- (iv) d_h is nondecreasing in α .
- (v) d_l is nondecreasing in $(1 - \beta)$.

Proof: See the Appendix.

□ **The arrival rate.** Consumers can obtain some information about potential firms to patronize from other consumers or other sources of information such as government or consumer organizations. Assume that in each period a times as many customers arrive at each high quality firm as at each low quality firm.

□ **Relative firm size.** Although many of the equilibrium quantities such as firm size are expected values, for ease of exposition, they will be spoken of as certain. Firms reach a constant size when their number of arrivals equals their number of departures. Define the following notation:

¹ See Smallwood and Conlisk (1979) for an analysis of a wide variety of assumptions about consumer flows and their implications for equilibrium firm size.

D_h = number of consumers departing from high quality firms;
 A_h = number of consumers arriving at high quality firms;
 G_h = number of high quality firms;
 N = number of consumers;
 N_h = number of consumers patronizing high quality firms; and
 X_h = size of high quality firm.

The same variables can be defined for low quality firms by replacing the subscript h with an l .

Consumer departures are determined by the departure rates,

$$D_h = d_h N_h \quad (2)$$

$$D_l = d_l N_l \quad (3)$$

Consumer arrivals are similarly determined by the relative arrival rate,

$$a = \frac{A_h / G_h}{A_l / G_l} \quad (4)$$

Setting arrivals equal to departures yields

$$\frac{X_h}{X_l} = a \frac{d_l}{d_h} \quad (5)$$

Let γ denote the number ad_l/d_h . The parameter γ is always greater than or equal to one. It is a composite measure of the extent to which arrivals and departures reward high quality firms relative to low quality firms.

3. Firm entry and product quality choice

■ Two simplifying assumptions are made. First, it is assumed that firms make a once-and-for-all quality choice upon entering the market. Although the subject of oscillation and variation in individual firms' quality is interesting, in many markets where quality is an issue, such as those for lawyers, mechanics, or opticians, changing quality may not be an important phenomenon. The costs associated with such changes may outweigh any benefits. In many cases the firm's quality depends on variables such as the quality of full-time personnel, capital equipment, and the professional's training, and none of these are easily variable on a short-term basis. Second, it is assumed that firms do not discount future income. If firms were concerned about the short-run income flows generated immediately upon entry as the firm grew to equilibrium size, there would be two changes. There would be less entry in general, and to the extent that the relative size advantage of being high quality is only manifested over time, there would be less incentive to produce high quality goods. But, these are changes only in degree; the qualitative properties of the model remain unchanged.

There are I firms and each firm's cost structure is some type $i \in [0, 1]$. A firm with cost structure i can produce high and low quality goods at a constant marginal cost of, respectively, c_h^i and c_l^i . Production of either type of good can only occur up to a capacity of k and involves fixed costs of $F + F_i$. Let $H(i)$ denote the distribution of firm types.

The number of firms choosing to be high and low quality is now easily determined. Each firm observes the size of high and low quality firms, X_h and X_l , and the market price, p . Taking these as given, each firm makes a profit maximizing quality and entry choice. The profits from producing a high quality good are

$$\Pi_h^i = X_h(p - c_h^i) - F - F_i \quad (6)$$

and those from producing a low quality good are

$$\Pi_i^l = X_l(p - c_i^l) - F - F_i. \quad (7)$$

For technical convenience it is assumed that the relationship between (p, X_h, X_l) and the number of high and low quality firms, G_h and G_l , is a function and is continuous in price. This embodies two different assumptions. First, the relationship is assumed to be a function instead of a correspondence to avoid the expositional difficulty of treating correspondences. The analysis remains unchanged if correspondences are allowed. The economic interpretation of this assumption is that each firm has at least a slightly different cost function from all other firms. Second, the continuity of the relationship corresponds to the assumption that each firm is infinitesimally small. As is usual in models with fixed costs and discrete firms, an equilibrium almost never exists precisely because of the integer problem; for demand to be met exactly firms will be producing at slightly over or slightly under capacity. The required deviation from capacity will be insignificant, however, in markets such as those for professional services where there is a large number of small firms.² To avoid notions of approximate equilibria, it is simply assumed that firms are infinitesimal. Denote the functions determining G_h and G_l by

$$G_h = g_h(p, X_h, X_l, F) \quad (8)$$

$$G_l = g_l(p, X_h, X_l, F). \quad (9)$$

(The parameters c_h^i , c_l^i , and F_i are omitted from (8) and (9) because they are fixed throughout the analysis.)

By rearranging (6) and (7) it can be seen that the firm will prefer to produce a *high* quality product if the market price is above the number p_i^* , defined by

$$p_i^* = c_h^i + \frac{c_h^i - c_l^i}{\frac{X_h}{X_l} - 1}. \quad (10)$$

In accord with Klein and Leffler (1981), call p_i^* the quality guaranteeing price for type i . The firm's behavior can be explained as follows. The advantage to producing high quality goods is that the firm obtains more customers; the disadvantage is that its costs rise. If price equaled c_h^i , there would be no advantage to generating more business by producing high quality goods; it would be more profitable to sell fewer goods, but make a profit on each. As price rises above c_h^i , the advantage to generating more business by producing high quality goods becomes more substantial until finally at some point it becomes large enough that producing high quality goods becomes the more profitable course of action.

4. Market equilibrium

- In equilibrium, high quality firms must be operating just at capacity:

$$X_h = k. \quad (11)$$

If all firms were operating at less than full capacity, downward pressure on price would occur (causing exit) until at least high quality firms were operating at capacity. At this point low quality firms would still be producing at less than capacity since they are $1/\gamma$ times the size of high quality firms. Low quality firms could not, however, lower prices unilaterally since they would thereby reveal themselves as being low quality. If a

² See Novshek (1980) for a formal proof of this.

price did result with demand for high quality firms exceeding capacity, access to the high quality firms would necessarily be rationed to maintain their size at capacity. In this situation, price would be driven upwards (causing entry) until high quality firms were just able to meet demand.

Condition (11) together with the condition that low quality firms be $1/\gamma$ times the size of high quality firms,

$$\frac{X_h}{X_l} = \gamma, \quad (12)$$

and the condition that industry output equal demand,

$$X_h g_h(p, X_h, X_l, F) + X_l g_l(p, X_h, X_l, F) = N, \quad (13)$$

determine equilibrium values for the three variables, X_h , X_l , and p . Other equilibrium values are then determined by substituting those three values into the following:

$$G_h = g_h(p, X_h, X_l, F) \quad (14)$$

$$G_l = g_l(p, X_h, X_l, F) \quad (15)$$

$$N_h = X_h g_h(p, X_h, X_l, F) \quad (16)$$

$$N_l = X_l g_l(p, X_h, X_l, F). \quad (17)$$

Proposition 2 gives sufficient conditions for existence of the equilibrium.

Proposition 2: Suppose that for large enough prices, industry output exceeds demand. That is, suppose,

$$kI > N. \quad (18)$$

Then an equilibrium price exists.

Proof: When price is less than the infimum of all c_i^j , industry output is zero. By (18), for large enough values of p , industry output exceeds N . Therefore, by the assumption about continuity of $g_h(\cdot)$ and $g_l(\cdot)$ in p , industry output must equal N at some price in between. *Q.E.D.*

Two points should be noted. First, if consumer search costs were high enough or the equilibrium fraction of high quality products was low enough, the (identical) consumers might simply prefer to purchase low quality products that were labelled as such for a lower price. This type of nonexistence of equilibrium is not formally analyzed in the model.

Second, the only type of nonuniqueness in equilibrium price that may occur is for there to be an interval of prices for which industry output is constant and equal to N . This would be a case in which increasing price in the interval caused no switchovers by consumers and no entry by firms. Thus, equilibrium values of G_h , G_l , N_h , and N_l would be uniquely determined even in this case. That is, equilibrium is always essentially unique. For expositional convenience the equilibrium price will be spoken of as being unique.

Proposition 3 describes the comparative statics of the equilibrium values.

Proposition 3:

- (i) N_h is nondecreasing and N_l is nonincreasing in γ and F .
- (ii) p is nondecreasing in F .

Proof: See the Appendix.

Recall that γ is a measure of the extent to which reputation operates in the market by rewarding high quality producers with more customers. Part (i) shows that the result of reputation's operating more effectively is to increase the fraction of high quality pro-

duction. From Proposition 1, improving the consumers' ability to judge product quality, either by lowering the probability (α) that high quality goods are mistaken as low quality or by lowering the probability (β) that low quality goods are mistaken as high quality, will alter departure rates in a way that increases γ and thus increases the fraction of high quality goods consumed. Similarly, policies which improve consumers' ability to differentiate between high and low quality firms in advance, thus increasing a , will increase γ and therefore increase the fraction of high quality goods produced. In particular, to the extent that advertising increases consumer information about the quality of various firms or about how to judge the quality of output, it should result in a greater fraction of high quality output.

More surprisingly, part (i) of Proposition 3 also shows that an increase in fixed costs causes the fraction of high quality products to increase. This can be illuminated by referring to part (ii). The increase in fixed costs causes equilibrium price to rise. This means that the quality-guaranteeing price of a larger number of firms is below the market price, and thus more high quality products are produced. This suggests that industry organizations may well promote product quality by creating large fixed costs for members in the form of training requirements or license fees, even if the training or licensing itself has no direct effect on quality. To the extent that advertising is a fixed cost, the same point applies: an increase in all firms' advertising expenditures may improve product quality, even if advertising communicates no useful information.

5. The consumer's choice of a departure rule

■ Previous sections were based on the assumption that consumers use a fixed departure rule in determining whether to continue patronizing a firm. This section analyzes the consumer's choice of departure rule and compares the equilibrium rule—the one that consumers will use—with the optimal rule—the one that maximizes consumers' welfare.

Two types of externalities might be expected to cause the equilibrium to depart from optimality. The first is the usual one; consumers' choice of how harsh a judge to be affects firms' incentives to produce high quality goods.³ Thus, a harsher decision rule benefits all consumers. The consumer, however, ignores this benefit when choosing a departure rule. The potential for this externality to appear is explicitly eliminated in this section by fixing the number of high and low quality firms.

We show that a second, more novel type of externality, which depends on the way in which information flows between consumers, exists. If the quality of a good or service depends on a number of poorly measurable factors, the most often communicated piece of information may simply be whether the consumer intends to continue to patronize his firm. If a consumer could simply report the vector describing the quality of each good he had purchased, the departure rule would be irrelevant to what information is transferred. But when the information flow between consumers consists principally of positive or negative recommendations—that is, whether the consumer will continue with a firm—instead of actual descriptions of the firm's performance, the consumer's choice of departure rule affects the information that he will communicate to other consumers. When choosing a departure rule, a consumer ignores the potential benefits to other consumers that accrue through recommendations.

This can be formally demonstrated as follows. Suppose that there are fixed numbers of high quality and low quality firms. In addition, restrict attention to symmetric equilibria—those in which all consumers use the same search rule. Consider a particular

³ This is equivalent to the externality that appears in models where consumers search for the lowest price. See, for example, Salop and Stiglitz (1977). See Farrell (1982b) for a similar treatment of the product quality issue.

consumer. Assume that all other consumers are using the departure rule r^* , and let r denote the particular consumer's choice of a departure rule. The particular consumer is then faced with the following optimization problem. First, he must choose a departure rule, r . Second, after departing from a firm he must choose a strategy for selecting a new firm. The consumer chooses a sequential strategy describing how to ask for recommendations and when finally to settle on a new firm. The nature and quality of information that these recommendations contain depend on the departure rule that other consumers use.⁴ The consumer can, therefore, rank potential departure rules for *other* consumers in terms of the usefulness of the statistical information they generate for him. Let the real-valued function δ represent this ranking. That is, $\delta(r^{*1}) > \delta(r^{*2})$ if r^{*1} provides a signal that the consumer prefers to that provided by r^{*2} .

The consumer's welfare has now been conceptualized as depending on two factors—the consumer's own choice of a departure rule, r , and the statistical quality of information provided by other consumers, δ . This latter variable in turn depends on other consumers' choice of a departure rule, r^* . This conceptualization of the consumers' welfare will be written as

$$W = W(r, \delta(r^*)), \quad (19)$$

where W is strictly increasing in δ . An equilibrium occurs when the consumer's choice of r , given r^* , is equal to r^* . That is, r_e is an *equilibrium rule* if and only if it satisfies

$$r_e \in \operatorname{argmax}_r W(r, \delta(r_e)). \quad (20)$$

An optimal rule is a rule which would maximize the typical consumer's welfare were everyone to use it. That is, r_o is an *optimal rule* if and only if it satisfies

$$r_o \in \operatorname{argmax}_r W(r, \delta(r)). \quad (21)$$

In general, the equilibrium and optimal rules will be different. Proposition 4 shows that the optimal rule produces at least as much useful information as the equilibrium rule does.

Proposition 4: Suppose that r_o and r_e are respectively optimal and equilibrium departure rules. Then $\delta(r_o) \geq \delta(r_e)$.

Proof: Suppose that $\delta(r_o) < \delta(r_e)$. Then $W(r_o, \delta(r_o)) < W(r_o, \delta(r_e))$. By the definition of r_e , $W(r_o, \delta(r_e)) \leq W(r_e, \delta(r_e))$. Together, these two inequalities imply that $W(r_o, \delta(r_o)) < W(r_e, \delta(r_e))$, thereby contradicting the fact that r_o is optimal. *Q.E.D.*

This result yields two conclusions. First, the source of this externality is the consumers' inability to transmit objective descriptions of firm performance to one another and their need to rely instead on positive or negative recommendations. Therefore, policies that facilitate the exchange of objective quality data among consumers might be welfare improving to the extent that they remove this externality. In the language of Beales, Craswell, and Salop (1981), government activity to establish a consumer metric may be warranted. Also, to the extent that advertising increases the information that consumers have and gives them a common language to describe products, it may facilitate the transfer of objective quality data among them. Second, this result highlights the more general

⁴ From Proposition 1, any rule r^* will generate some departure rates, d_h^* and d_l^* . Then a recommendation is simply a signal with the following likelihoods. Let + and - denote a positive and a negative recommendation and h and l denote whether the firm in question is really high or low quality:

$$\Pr(+/h) = (1 - d_h^*)$$

$$\Pr(+/l) = (1 - d_l^*).$$

point that the specific manner in which information is transferred among consumers may have a significant impact on the market outcome. This suggests that analysis of the manner in which information flows between consumers and its consequences for the market would be a fruitful area for further research.

Appendix

■ Proofs of Propositions 1 and 3 follow.

Proposition 1:

(i) On average, $B'_h n$ customers have been with the firm t periods before. Therefore, the firm's total number of customers is

$$\left[\sum_{t=1}^{\infty} B'_h n \right] + n. \tag{A1}$$

The result in (i) follows from the definition in (1).

(ii) If n customers arrive every period and the expected number of customers stays constant at n/d_h , then the expected number of customers that leave each period is also n . That is, on average d_h of the firm's total customers leave each period. The case for a low quality firm is similar.

(iii) It is clearly sufficient to show that $B'_h \geq B'_l$ for every t . This follows from the fact that the probability of a high quality observation is greater when the consumer patronizes a high quality firm.

(iv) and (v) It is sufficient to show that each B'_h is nonincreasing in α and B'_l is nonincreasing in $(1 - \beta)$. This is straightforward. *Q.E.D.*

Proposition 3: Since $N_h + N_l = N$, it suffices to prove the proposition for N_h . Substitution of (11) and (12) into (13) yields

$$k f_h(p, \gamma, F) + \gamma k f_l(p, \gamma, F) = N, \tag{A2}$$

where

$$f_h(p, \gamma, F) = g_h(p, k, \gamma k, F) \tag{A3}$$

$$f_l(p, \gamma, F) = g_l(p, k, \gamma k, F). \tag{A4}$$

Using subscripts to denote partial derivatives of f_h and f_l , total differentiation of (A2) yields

$$\frac{\partial N_h}{\partial \gamma} = \left[\frac{k}{\gamma f_{h\gamma} + f_{l\gamma}} \right] \left[(f_{h\gamma} f_{lp} - f_{l\gamma} f_{hp}) + \frac{f_l f_{hp}}{\gamma} \right] \tag{A5}$$

$$\frac{\partial N_h}{\partial F} = \left[\frac{k}{\gamma f_{hF} + f_{lF}} \right] [f_{hF} f_{lp} - f_{lF} f_{hp}] \tag{A6}$$

$$\frac{\partial p}{\partial F} = \left[\frac{-1}{\gamma f_{hp} + f_{lp}} \right] \left[f_{hF} + \frac{f_{lF}}{\gamma} \right]. \tag{A7}$$

A number of terms need to be signed to prove the proposition. These will be listed and proved as lemmas. Lemma 1 merely collects the properties of f_h and f_l for later reference.

Lemma 1:

- (i) f_h is nondecreasing in $\{p, \gamma, k\}$;
- (ii) f_h is nonincreasing in $\{F\}$;
- (iii) f_l is nondecreasing in $\{k\}$;
- (iv) f_l is nonincreasing in $\{F, \gamma\}$;

(v) $(f_h + f_l)$ is nondecreasing in $\{k, p\}$; and

(vi) $(f_h + f_l)$ is nonincreasing in $\{\gamma, F\}$.

Proof: The proof is straightforward and available from the author upon request. *Q.E.D.*

Lemma 2:

$$\gamma f_{hp} + f_{lp} \geq 0. \quad (\text{A8})$$

Proof: By Lemma 1, $f_{hp} \geq 0$ and $f_{hp} + f_{lp} \geq 0$. The result now follows immediately because $\gamma \geq 1$. *Q.E.D.*

Lemma 3: $f_{h\gamma}f_{lp} - f_{l\gamma}f_{hp} \geq 0$.

Proof: By Lemma 1, $f_{hp} \geq 0$ and $f_{hp} + f_{lp} \geq 0$. Therefore, either $|f_{hp}| \geq |f_{lp}|$ or $f_{lp} \geq 0$.

By Lemma 1, $f_{h\gamma} \geq 0$, $f_{l\gamma} \geq 0$, and $f_{h\gamma} + f_{l\gamma} \leq 0$. Therefore, $|f_{l\gamma}| \geq |f_{h\gamma}|$. Suppose that $|f_{hp}| \geq |f_{lp}|$. Then,

$$\left. \begin{array}{l} |f_{hp}| \geq |f_{lp}| \\ |f_{l\gamma}| \geq |f_{h\gamma}| \end{array} \right\} \Rightarrow |f_{hp}f_{l\gamma}| \geq |f_{lp}f_{h\gamma}|. \quad (\text{A9})$$

Since $-f_{hp}f_{l\gamma} \geq 0$, Lemma 3 is true. Now suppose that $f_{lp} \geq 0$. Then Lemma 3 is clearly true because both terms in the sum are nonnegative. *Q.E.D.*

Lemma 4: $f_{hF}f_{lp} - f_{lF}f_{hp} \geq 0$.

Proof: The proof is similar to that of Lemma 3. *Q.E.D.*

Proposition 3 now follows directly. *Q.E.D.*

References

- AKERLOF, G.A. "The Market for 'Lemons': Quality Uncertainty and the Market Mechanism." *Quarterly Journal of Economics*, Vol. 84 (August 1970), pp. 488-500.
- BEALES, H., CRASWELL, R., AND SALOP, S. "Information Remedies for Consumer Protection." *American Economic Review*, Vol. 71, No. 2 (May 1981), pp. 410-413.
- BENHAM, L. AND BENHAM, A. "Regulating the Professions: A Perspective on Information Control." *Journal of Law and Economics*, Vol. 18 (1975), pp. 421-447.
- DARBY, M. AND KARNI, E. "Free Competition and the Optimal Amount of Fraud." *Journal of Law and Economics*, Vol. 16 (April 1973), pp. 67-88.
- DYBVIK, P.H. AND SPATT, C.S. "Does It Pay to Maintain a Reputation?" Mimeo, March 1982.
- FARRELL, J. "Introductory Offers and Entry with Rational (but Uninformed) Buyers." Mimeo, March 1982a.
- . "Free Entry and Responsibility." Mimeo, May 1982b.
- KLEIN, B. AND LEFFLER, K.B. "The Role of Market Forces in Assuring Contractual Performance." *Journal of Political Economy*, Vol. 89 (August 1981), pp. 615-641.
- NOVSHEK, W. "Cournot Equilibrium with Free Entry." *Review of Economic Studies*, Vol. 47 (1980), pp. 473-486.
- ROGERSON, W.P. "Reputation and Product Quality." IMSSS Technical Report No. 384, 1982.
- SHAPIRO, C. "Premiums for High Quality Products as Rents to Reputation." Woodrow Wilson School Discussion Paper #6, April 1981.
- . "Consumer Information, Product Quality and Seller Reputation." *Bell Journal of Economics*, Vol. 13, No. 1 (Spring 1982), pp. 20-35.
- SMALLWOOD, D. AND CONLISK, J. "Product Quality in Markets Where Consumers Are Imperfectly Informed." *Quarterly Journal of Economics*, Vol. 93 (February 1979), pp. 1-23.
- VON WEIZSÄCKER, C.C. "A Welfare Analysis of Barriers to Entry." *Bell Journal of Economics*, Vol. 11, No. 2 (Autumn 1980), pp. 399-420.

89

Beyond Disclosure: The Case for Banning Contingent Commissions

Daniel Schwarcz[†]

Introduction.....	290
I. The Contingent Commission Debate: Increased Steering vs. Reduced Adverse Selection	
A. <i>Contingent Commissions and Inefficient Steering</i>	296
B. <i>Contingent Commissions and Adverse Selection</i>	303
C. <i>Other Potential Benefits of Contingent Commissions</i>	305
1. <i>Contingent Commissions as a Method of Expanding Insurance Coverage</i>	306
2. <i>Contingent Commissions as a Protector of Small Agencies</i>	308
3. <i>Contingent Commissions as a Protector of Small Insurers</i>	309
4. <i>Contingent Commissions as a Facilitator of Limited Insurer-Intermediary Interactions</i>	310
II. Contingent Commissions in Consumer Insurance Markets	
A. <i>Disclosure, Steering, and Trilateral Dilemmas</i>	311
B. <i>Evaluating the Adverse Selection Argument</i>	312
C. <i>Toward Banning Contingent Commissions</i>	319
III. Contingent Commissions in Insurance Markets with Sophisticated Insureds	
A. <i>Steering and Signaling in Sophisticated Insurance Markets</i>	325
B. <i>Mimicry: A New Adverse Selection Model for Sophisticated Insurance Markets</i>	328
1. <i>The Mimicry Assumption</i>	329
2. <i>The Consequences of Mimicry</i>	330

[†] Climenko Fellow and Lecturer in Law, Harvard Law School. Associate Professor of Law Designate, University of Minnesota Law School. Harvard Law School, J.D.; Amherst College, A.B. For helpful comments and discussion, I thank Tom Baker, Glenn Cohen, Neil Doherty, Sean Fitzpatrick, Sydney Foster, Jack Goldsmith, Carissa Byrne Hessick, Bert Huang, Howell Jackson, Shannon McCormick, J.J. Prescott, Steven Schwarcz, Steven Shavell, Peter Siegelman, Matthew Stephenson, and Sean Williams. This Article benefited from comments received in workshops at Harvard Law School and the New England Insurance Law Reading Group. Harold Birnbaum and Daniel Epps provided excellent research assistance.

C. <i>A Possible Need for Intervention in Sophisticated Insurance Markets</i>	335
Conclusion	335

INTRODUCTION

In November 2004, New York's prominent Attorney General (and now Governor) Eliot Spitzer accused insurance broker Marsh & McLennan of defrauding its clients by accepting "contingent commissions."¹ Contingent commissions are bonuses that insurers pay to brokers and independent agents (collectively "independent intermediaries" or "producers") for bringing the insurer a particularly large volume of profitable customers.² Since Spitzer's initial attack on contingent commissions, his office has parlayed similar accusations³ into six settlement agreements with major insurance industry companies, twenty guilty pleas from these companies' executives and officers, and approximately \$3 billion in restitution and penalties.⁴ According to the latest count, more than twenty states have opened their own investigations into misconduct in the insurance industry.⁵ And seven states, with more expected to follow, have passed legislation to combat the types of practices Spitzer uncovered.⁶ Federal officials have monitored these responses closely, with some suggesting that they expose fundamental problems with lodging insurance regulation at the state—rather than federal—level.⁷ In short, during the last two years contingent commissions have caused a scandal unparalleled

1. Press Release, Office of the N.Y. State Att'y Gen., Investigation Reveals Widespread Corruption in Insurance Industry (Oct. 14, 2004), available at http://www.oag.state.ny.us/press/2004/oct/oct14a_04.html.

2. See J. David Cummins & Neil A. Doherty, *The Economics of Insurance Intermediaries*, 73 J. RISK & INS. 359, 375 (2006).

3. The prominence of Spitzer's initial attack on contingent commissions is partially attributable to the fact that employees at Marsh & McClellan had solicited purposefully inflated price quotations from some insurers. This "bid rigging" allowed Marsh to steer its clients to particular insurers by falsely convincing them that the insurer had prevailed in a competitive process of price quotation. The sole benefit to Marsh, and its guilty employees, was the receipt of additional contingent commissions. See Sean M. Fitzpatrick, *The Small Laws: Eliot Spitzer and the Way to Insurance Market Reform*, 74 FORDHAM L. REV. 3041, 3046-47 (2006). Subsequent investigations, however, have revealed that Marsh's undeniably fraudulent practice of bid-rigging was isolated to a small number of the giant insurance broker's employees; bid-rigging was not, as initially seemed plausible, widespread among insurance brokers and insurers in the property-casualty industry. See *id.*

4. Press Release, Office of the N.Y. State Att'y Gen., Ace Settles Bid-Rigging Probe; Agreement Is Part of Ongoing Effort To Restore Competition in Insurance Industry (Apr. 26, 2006), available at http://www.oag.state.ny.us/press/2006/apr/apr26a_06.html.

5. Fitzpatrick, *supra* note 3, at 3050 n.39.

6. *Id.* at 3064.

7. See R.J. Lehmann, *Specter Grills Witnesses on Bid Rigging Prosecutions*, BESTWIRE, June 21, 2006.

Beyond Disclosure: The Case for Banning Contingent Commissions

in the history of the property-casualty insurance industry.⁸

According to critics of the practice, contingent commissions create a conflict of interest for ostensibly independent intermediaries. Unlike ordinary premium-based commissions, which are a fixed percentage of an individual insurance customer's premiums, contingent commissions are paid to intermediaries based on the volume and profitability of business they refer to insurers.⁹ The size and structure of the contingent commissions that insurers offer to intermediaries therefore vary significantly. This variability means that intermediaries can often increase their commissions by "steering" clients to insurers that provide suboptimal coverage for their customers' needs.¹⁰

Despite the recent controversy surrounding contingent commissions, these payments and similar performance-based commissions remain common in most property-casualty insurance markets.¹¹ Contingent commissions are ubiquitous among independent insurance agents, who provide 40% of ordinary consumers with insurance products covering their homes, automobiles, or small businesses.¹² Contingent commissions are also widely used in insurance markets involving sophisticated insurance purchasers—where brokers rather than agents are the dominant type of intermediary.¹³ Although each of the four largest insurance brokers have disclaimed contingent commissions in response

8. See Fitzpatrick, *supra* note 3, at 3041.

9. See Cummins & Doherty, *supra* note 2, at 379.

10. Complaint at 7-8, *State v. Marsh & McLennan Cos.*, No. 04403342 (N.Y. Sup. Ct. Oct. 14, 2004), available at http://www.oag.state.ny.us/press/2004/oct/oct14a_04_attach1.pdf [hereinafter *Marsh Complaint*].

11. See J. ROBERT HUNTER, CONSUMER FED'N OF AM., CONTINGENT INSURANCE COMMISSIONS: IMPLICATIONS FOR CONSUMERS 2 (2005), http://www.consumerfed.org/pdfs/contingent_commissions_study.PDF.

12. *Id.* at 4 ("[C]ontingent arrangements exist in most lines of insurance sold to consumers In personal auto and homeowners insurance, agent-based insurers cover about 40 percent of the market."). Captive agents or direct writers, both of whom sell only a single insurer's policies, obviously do not receive contingent commissions.

13. In theory, independent agents and brokers differ in the extent to which they are legal agents of the policyholder or the insurer. See KENNETH S. ABRAHAM, *INSURANCE LAW & REGULATION* 56-57 (3d ed. 2000); 7 ERIC MILLS HOLMES, *HOLMES' APPLEMAN ON INSURANCE 2D: LAW OF INSURANCE AGENTS* § 47.5, at 326 (1998). Insurance brokers are typically described as general legal agents of the insurance consumer who are free to place the consumer's business with whatever insurer the consumer or broker chooses, whereas independent insurance agents ostensibly function as independent contractors who sell the products of multiple different insurers. See ABRAHAM, *supra*, at 56-57. In practice, the difference between these two types of intermediaries is often hazy at best. See TOM BAKER, *INSURANCE LAW AND POLICY* 66 (2003); Colin Sammon, Comment, *Insurance Agent and Broker Liability: Crossing the Two Way Street*, 29 OHIO N.U. L. REV. 237, 238 (2002). Indeed, several states have abandoned the distinction between agents and brokers altogether. See HOLMES, *supra*, § 47.6, at 338; Fitzpatrick, *supra* note 3, at 3054 (noting Connecticut as one example). To the extent that there is a meaningful practical distinction between these two intermediaries, it is that brokers tend to serve relatively sophisticated clients, whereas agents tend to serve less sophisticated ordinary consumers. See Fitzpatrick, *supra* note 3, at 3055. It is this latter distinction—based on the level of sophistication of the insurance purchaser—that is relevant for the purposes of policy analysis. Consequently, this Article structures its analysis around consumer sophistication rather than the agent/broker distinction.

to pressure from Spitzer's office,¹⁴ the majority of brokers have refused to do so.¹⁵ And a handful of insurers that currently claim to have abandoned contingent commissions have merely adopted "supplemental compensation" arrangements, which retain the same performance-based contingency structure that generates the underlying conflict of interest.¹⁶ In effect, such "supplemental compensation arrangements" are contingent commissions by a different name.¹⁷

Independent insurance intermediaries defend their continued receipt of contingent commissions by arguing that competition can adequately address the payments' potential dangers.¹⁸ To date, this argument has apparently convinced state regulators and legislatures: the vast majority of reform proposals permit contingent commissions so long as they are adequately disclosed.¹⁹ And each of the approximately half-dozen states that has enacted reforms has merely imposed disclosure requirements.²⁰ Indeed, Spitzer himself has suggested that contingent commissions might be acceptable in some parts of the industry if

14. Aside from Marsh, these are Willis Group Holdings Ltd., Aon Corporation, and Arthur J. Gallagher & Co. See Cummins & Doherty, *supra* note 2, at 365 tbl.2; see also Press Release, Arthur J. Gallagher & Co., Arthur J. Gallagher & Co. Announces Subpoena from Connecticut Attorney General, Elimination of Volume & Profit Based Contingent Commissions, and Third Quarter 2004 Financial Results (Oct. 26, 2004), available at http://media.corporateir.net/media_files/irol/10/104111/AJG_Q3_2004_Earnings_Release.pdf; Press Release, Office of the N.Y. State Att'y Gen., Aon Settles Corruption Probe: Leading Insurance Broker Agrees To Pay \$190 Million and Adopt Sweeping Reforms (Mar. 4, 2005), available at http://www.oag.state.ny.us/press/2005/mar/mar04a_05.html; Press Release, Office of the N.Y. State Att'y Gen., Insurance Broker Agrees to Sweeping Reforms: Marsh To Pay \$850 Million in Restitution and Ban Contingent Commissions (Jan. 31, 2005), available at http://www.oag.state.ny.us/press/2005/jan/marshsettlement_pr.pdf; Press Release, Willis Group Holdings, Willis Group Holdings To End Practice of Contingency Agreements with Insurance Carriers (Oct. 21, 2004), available at http://www.willis.com/news/news_attachments/end_Contingency.pdf.

15. See David Dwanka, *Mid-Level Insurance Brokers Defend Contingent Commissions Amid Growing Criticism*, BESTWIRE, May 8, 2006.

16. See Sally Roberts, *Compensation Shake-Up Continues; Chubb Pays \$17M, Ends All Contingents*, BUS. INS., Dec. 25, 2006, at 1; Press Release, Office of the N.Y. State Att'y Gen., Nation's Largest Disability Insurer Agrees to Sweeping Reforms: UnumProvident To Eliminate Contingent Commissions; Fully Disclose Broker Compensation (Nov. 1, 2006), available at http://www.oag.state.ny.us/press/2006/nov/nov01a_06.html.

17. Accordingly, this Article uses the term "contingent commission" to encompass "supplemental compensation" agreements. Cf. Roberts, *supra* note 16, at 1 (noting that the Chubb agreement states that "a fixed commission paid to a producer, set prior to the sale of a particular insurance product, and that may be based on, among other things, the prior year's performance of the producer" is not considered a contingent compensation"). For more on the distinctive issues that supplemental compensation arrangements pose, see *infra* notes 163-164 and accompanying text.

18. See Dwanka, *supra* note 15.

19. See Fitzpatrick, *supra* note 3, at 3050 n.39 (noting that over twenty states have opened investigations into contingent commissions as of May 2006, and that "[l]egislative and administrative market reform efforts have targeted disclosure of compensation received by insurance producers, as opposed to proscribing contingent commissions themselves"). These disclosure requirements do not require disclosing the size of the contingent commissions associated with a particular transaction. Indeed, such disclosure might well be impossible given that contingent commissions are calculated at year's end.

20. See *id.* at 3064. Many of the most important states for insurers, including California and New York, have not yet adopted reforms. *Id.*

Beyond Disclosure: The Case for Banning Contingent Commissions

insurers informed consumers about their existence.²¹

Significantly, academic commentary has supported this disclosure-based response to contingent commission schemes. The most notable claim in the extant literature is that because contingent commissions are based on insurers' profitability, they encourage producers to inform insurers about the risk characteristics of potential purchasers ("insureds") in ways that would otherwise be impossible.²² Some have therefore concluded that contingent commissions can help consumers by reducing information asymmetries between them and their insurers. This would limit adverse selection, a highly theorized problem in insurance markets, which occurs when high-risk insureds purchase more insurance than low-risk insureds, potentially causing premiums to spiral upward as "good" risks forego insurance altogether. Given this and other potential benefits of contingent commissions, commentators have argued that insurance purchasers, so long as they are informed, should be allowed to choose brokers and agents who accept contingent commissions.²³ According to these writers, not only would meaningful disclosure empower insureds over regulators, but it also would prompt more careful scrutiny of intermediaries' actions, thereby limiting the likelihood that insureds would be steered toward suboptimal insurance.

This Article questions the emerging consensus among state officials and academics that disclosure is a sufficient solution to the problems that contingent commissions pose. Both the substance and the strength of this challenge depend upon the sophistication and knowledge of insurance purchasers. For this reason, this Article first analyzes consumer insurance markets—such as homeowners, renters, and automobile insurance—where purchasers tend to be relatively unfamiliar with many elements of the insurance industry. It then proceeds to commercial insurance markets, such as directors' and officers' (D&O) insurance and business automobile insurance, where purchasers are assumed to be knowledgeable and rational about their insurance options.

The Article concludes that a disclosure requirement in consumer insurance

21. In a speech to the National Press Club, Spitzer stated that the use of contingent commissions "may be appropriate" in certain segments of the insurance industry when they are disclosed to consumers. Press Release, Nat'l Ass'n of Prof'l Ins. Agents, PIA National Encouraged by Remarks Made by N.Y. Attorney General Eliot Spitzer on Contingent Commissions (Jan. 31, 2005), available at <http://www.pianet.com/NewsCenter/PressReleases/1-31-05.htm>.

22. See Cummins & Doherty, *supra* note 2, at 386-89; Fitzpatrick, *supra* note 3, at 3060-61; see also Gary Biglaiser, *Middlemen as Experts*, 24 RAND J. ECON. 212 (1993) (showing that middlemen can help reduce adverse selection in a variety of markets, but not discussing the role of a contingent commission payment structure in this process).

23. See Fitzpatrick, *supra* note 3, at 3066-71; see also Cummins & Doherty, *supra* note 2, at 394 (suggesting that contingent commissions should not be banned because they can be beneficial to clients); Neil A. Doherty & Alexander Muermann, *Insuring the Uninsurable: Brokers and Incomplete Insurance Contracts* 18 (Ctr. for Fin. Studies, Working Paper No. 24, 2005), available at http://www.ifk-cfs.de/papers/05_24.pdf.

markets is unlikely to address meaningfully the core risk of contingent commissions—the potential for inefficient steering.²⁴ In fact, contingent commissions are merely one example of a common type of regulatory problem, termed a “trilateral dilemma” by Professor Howell Jackson, which lawmakers consistently regulate in ways that go beyond mere disclosure.²⁵ In a generic trilateral dilemma, market intermediaries extract side payments from other professionals for steering business to them.²⁶ For instance, a real estate agent may recommend a particular lawyer to his customers in exchange for kickbacks from the lawyer.²⁷ Trilateral dilemmas tend to be immune to disclosure-based remedies for three basic reasons, each of which squarely applies to contingent commissions. First, consumers who rely on intermediaries to recommend other service providers generally have an inherently limited capacity to assess the end-service provider’s relative strengths and weaknesses. Second, because intermediaries interact closely with their customers, they can discriminate between sophisticated and unsophisticated customers, taking advantage of the latter by giving them biased advice while providing more objective advice to the former. Finally, customers often have a relationship of trust with their intermediary that blunts any tendency to scrutinize the advice they receive. Just as each of these considerations has generally supported robust government regulation of most trilateral dilemmas, they also support affirmative government intervention in the case of contingent commissions.

Not only are the risks of contingent commissions large in consumer insurance markets, but the benefits of these payments are minimal at best. The core benefit of reduced adverse selection that commentators have offered in support of contingent commissions is both significantly overstated and highly speculative in consumer insurance markets. Extensive empirical research has consistently shown that adverse selection is not a significant problem in these insurance markets.²⁸ Moreover, there is little empirical evidence that contingent commissions really do improve insurers’ information about potential insureds, and significant theoretical reasons to believe that they do not. Weighing these costs and benefits of contingent commissions, the Article suggests that the optimal regulatory solution may be to ban contingent commission arrangements in consumer insurance markets.

A disclosure-based response to contingent commissions in commercial insurance markets, where purchasers tend to be sophisticated about insurance

24. See *infra* Section II.A.

25. Howell E. Jackson & Laurie Burlingame, *Kickbacks or Compensation: The Case of Yield Spread Premiums*, 12 STAN. J.L. BUS. & FIN. (forthcoming 2007); see also Howell Jackson, *The Trilateral Dilemma* (July 2006) (unpublished manuscript, on file with author).

26. See Jackson & Burlingame, *supra* 25.

27. See *id.*

28. See *infra* Section II.B. See generally Peter Siegelman, *Adverse Selection in Insurance Markets: An Exaggerated Threat*, 113 YALE L.J. 1223 (2004).

Beyond Disclosure: The Case for Banning Contingent Commissions

options, is more likely to be effective than in consumer insurance markets. But this Article demonstrates that even when insurance purchasers are both informed and rational, there are still reasons to be concerned about a disclosure-based response. First, disclosure of contingent commissions may result in rational and informed insurance purchasers seeking out intermediaries who accept contingent commissions even though they would prefer such commissions to be banned due to the risk of steering. The reason for this divergence between market behavior and individual purchaser preferences is that those who purchase insurance through an intermediary that does not receive contingent commissions may signal to insurers that they view themselves to be high-risk. If so, then insurers will charge an increased premium for consumers who purchase insurance in this way, undermining the extent to which consumers' informed market choices reflect their actual preferences regarding contingent commissions. Second, the adverse selection defense of contingent commissions is also questionable in commercial insurance markets. Contingent commissions may erode the relationship of trust between independent intermediaries and their clients and consequently shift, rather than eliminate, the information asymmetry that results in adverse selection.

Part I of this Article reviews the arguments in the extant literature about the costs and benefits of contingent commissions. It shows that the key concern with respect to contingent commissions is the risk associated with inefficient steering. It also explains how contingent commissions may improve insurance markets by reducing adverse selection. Part II applies the framework developed in Part I to consumer insurance markets. It argues that the limitations of disclosure in solving trilateral dilemmas create a significant risk that insurance intermediaries will continue to steer their customers toward inefficient arrangements in the absence of more robust regulation. It also argues that the risks associated with adverse selection in consumer property-casualty insurance markets are small, and that they are not appreciably reduced by contingent commission payments. For these reasons, Part II concludes that lawmakers should consider banning contingent commission payments in consumer insurance markets. Part III applies the framework developed in Part I to commercial insurance markets. It develops a theoretical model that shows that even fully informed insurance purchasers may not make decisions about their insurance intermediaries that reflect their actual preferences regarding contingent commissions. It also shows that the capacity of contingent commissions to reduce adverse selection is limited because of their potential to induce strategic, inefficient behavior by rational high-risk insureds. Part III concludes that disclosure in commercial insurance markets, like disclosure in consumer markets, may be an insufficient regulatory solution, but that empirical work is needed before a policy prescription can be endorsed.

I. THE CONTINGENT COMMISSION DEBATE: INCREASED STEERING VS.
REDUCED ADVERSE SELECTION

The current debate about contingent commissions focuses on their impact on the incentives of independent intermediaries. These payments shift such incentives to be less aligned with those of insurance consumers and more aligned with those of insurers. As a result, independent intermediaries that receive contingent commissions may tend to act in ways that initially appear less favorable for the consumer and more favorable for the insurer.

This Part reviews, but does not critique, the two key ways in which the literature has suggested this may occur. Section A describes the argument that contingent commissions cause intermediaries to steer their customers to suboptimal insurance arrangements. Section B then reviews the primary defense of contingent commission payments: that they create an incentive for intermediaries to share relevant customer information with insurers in a way that limits adverse selection. Section C considers some alternative defenses of contingent commissions, but concludes that they are generally unpersuasive. This Part therefore seeks to frame the basic contingent commission debate as involving two competing claims: that contingent commissions increase steering, and that they decrease adverse selection. The remainder of the Article builds off of this framework.

A. *Contingent Commissions and Inefficient Steering*

Independent insurance intermediaries help facilitate efficient purchases by matching customers with insurance options that fit their needs and preferences. Insurers vary substantially based on their reputations for claims handling,²⁹ financial strength,³⁰ and risk management services.³¹ Even the scope of coverage that an insurer provides—in both its base policies and its endorsements—can vary.³² The role of independent intermediaries is to help their customers understand these variations in quality and assess what policy

29. See *Consumer Reports Investigates: Surviving the 'Hard Market' in Homeowners Insurance*, CONSUMER REP., Sept. 2004, at 36 (describing how homeowner insurers differ significantly with regard to the percentage of customers reporting problems with claims or delayed payment); *Homeowners Insurance: Report*, CONSUMER REP., Jan. 1999, at 16 (suggesting that consumers, when choosing an insurer, consider customer satisfaction with claims handling among other factors).

30. See generally A.M. Best, Guide to Best's Financial Strength Ratings (Jan. 2, 2007), <http://www.ambest.com/ratings/guide.asp>.

31. Some homeowners insurance companies, for instance, provide their insureds with free washing machine hoses in order to limit the risk of flooding. See Meg Green, *Top of Their Game*, BEST'S REV., Dec. 2006, at 32.

32. See Daniel Schwarcz, *A Products Liability Theory for the Judicial Regulation of Insurance Policies*, 48 WM. & MARY L. REV. 1389, 1417-18 (2007) (noting that large insurers can deviate from ISO standard form contracts because they can generate sufficient actuarial data from their own pool of insureds).

Beyond Disclosure: The Case for Banning Contingent Commissions

type is best for their needs.³³ Additionally, independent intermediaries can help their customers identify pricing differentials among insurers that are not attributable to differences in quality. This is not uncommon even in price-competitive insurance markets: insurers often differ in the focus of their underwriting, meaning that they are willing to offer certain types of individuals—for instance, those who do not drink any alcohol—lower rates than other insurers.³⁴ Independent intermediaries can help match their customers with the insurer that employs underwriting criteria most favorable for the customer.

Contingent commissions may cause ostensibly independent intermediaries to deviate from this market-matching role and to steer insureds to suboptimal insurance arrangements.³⁵ To understand why, consider the basic structure of contingent commissions. These payments are bonuses, which insurers pay to intermediaries in addition to ordinary premium commissions. While premium commissions are stable—a simple percentage of an insured's total premium—contingent commission arrangements vary considerably.³⁶ In general, the size of an intermediary's contingent commission is based on two variables: (1) the amount of insurance business that a particular intermediary refers to the insurer, as measured in total premiums; and (2) the profitability of that business, which is usually measured by the insurer's loss ratio.³⁷ In most cases, intermediaries are only entitled to contingent commissions if they meet threshold levels of both sales volume and profitability.³⁸ Once intermediaries reach these qualifying levels, their commissions typically increase with better results along either dimension.³⁹

This payment structure is likely to alter the ways in which independent intermediaries direct their customers to insurers. Most obviously, independent intermediaries may steer their clients to insurers that pay contingent

33. See ABRAHAM, *supra* note 13, at 56-66; see also Cummins & Doherty, *supra* note 2, at 360.

34. See generally Green, *supra* note 31, at 26 (describing how some of the most profitable property-casualty insurers focus on underwriting only particularly safe risks, and pass off some of the resulting cost savings to their insureds).

35. See HUNTER, *supra* note 11, at 2-8; Press Release, *supra* note 1.

36. See Cummins & Doherty, *supra* note 2, at 374-75.

37. See Jeffrey Wilder, *Competing for the Effort of a Common Agent: Contingency Fees in Commercial Insurance* 5 (U.S. Dep't of Justice, Antitrust Div., Econ. Analysis Group Working Paper No. EAG03-4, 2004), available at <http://ssrn.com/abstract=418061>. The loss ratio is the "ratio between premiums paid and losses incurred during a given period." BLACK'S LAW DICTIONARY 958 (7th ed. 1999). Premiums on both new policies and policy renewals are generally treated similarly in these calculations, which are almost always made on a yearly basis. Wilder, *supra*, at 5.

38. See Wilder, *supra* note 37, at 5. In some cases, contingent commission arrangements may be based only on profitability, not volume. However, according to Cummins and Doherty, "the great majority of the arrangements covering the smaller intermediaries is based on the profitability of the business written or profitability and volume." Cummins & Doherty, *supra* note 2, at 379. Marsh and McLennan's contingent commission arrangements were distinctive in this regard, paying Marsh solely based on the volume of premiums that it brought to a given insurer. See *id.* at 16 n.20.

39. See Wilder, *supra* note 37, at 5.

commissions over those that do not: intermediaries will generally earn more in contingent commissions when their customers buy insurance through an insurer with whom the intermediary has a contingent commission contract.⁴⁰ Contingent commissions may also induce intermediaries to steer their customers to certain insurers in a number of less obvious situations. For instance, even when an independent intermediary can earn contingent commissions from multiple different insurers, the intermediary may maximize its profits by steering its customers to an insurer that offers the most lucrative commission. Alternatively, an intermediary might profit by steering a customer to an insurer whose contingent commission contract has a minimum-volume requirement on the cusp of being satisfied, even though a different insurer was the best match for that particular consumer.⁴¹ Yet another possibility is that an intermediary might profit by steering relatively high-risk insureds to insurers who do not offer contingent commissions, so as to maintain a good “loss ratio” with the insurers who do offer contingent commissions.⁴²

To illustrate these dynamics more concretely, consider an insurance agent who has contingent commission arrangements with two insurers, *X* and *Y*. The arrangement with insurer *X* provides that the agent is entitled to an additional 1% of the insured’s premium if the agent’s book of business—the customers the agent successfully refers to the insurer—reaches \$1 million of premiums for the year and generates \$100,000 of profit. The arrangement with *Y* is identical, except that it entitles the intermediary to an additional 2% of the insured’s premium if the triggering conditions are met. Initially the agent will have an incentive to steer customers to both *X* and *Y*, rather than to other insurers, so that he can meet the million-dollar premium requirement for both insurers. If he reaches this plateau with *Y* but not before year’s end, he may then start

40. One exception, noted below, is that an intermediary may actually maximize contingent commissions by steering an insured who is high-risk to an insurer with whom the intermediary does not have a contingent commission arrangement.

41. See Wilder, *supra* note 37, at 19. Another potential risk for consumers is that contingent commissions may cause independent intermediaries to be insufficiently aggressive in seeking payment for their client’s insurance claim. See HUNTER, *supra* note 11, at 7-8. Generally independent intermediaries not only help their customers select appropriate coverage but also manage the ongoing relationship between the customer and insurer, transmitting premium payments and negotiating claims. See generally ROBERT H. JERRY II, UNDERSTANDING INSURANCE LAW § 35 (1987) (describing the role of intermediaries in facilitating the purchase of insurance). Because contingent commissions reward independent intermediaries based on the insurer’s profitability, they create a disincentive for intermediaries to maximize the claims payment their clients receive: the less the client receives, the more profitable the intermediaries’ book of business is with the insurer, and thus the more the intermediary stands to gain in contingent commissions. The danger of this effect is unclear, however, as insurance consumers are likely to be quite sensitive to their intermediaries’ capacity to successfully negotiate claims; payment on claims is likely to be almost as salient a consideration for most consumers as policy price. See Fitzpatrick, *supra* note 3, at 3062 (“Anyone with practical experiences in the insurance business knows that customers make lasting judgments about intermediaries and insurers based on their behavior in the claims-paying context . . .”).

42. This last scenario may be an example of efficient steering. For elaboration, see *infra* text accompanying notes 68-71.

Beyond Disclosure: The Case for Banning Contingent Commissions

referring more customers to *X*. Once the agent is comfortable that he will meet the premium requirements for both *X* and *Y*, he may then start referring lower-risk customers to *Y* in order to maximize *Y*'s profits, and thus to increase the likelihood of hitting the profit target that is linked to the payment of the contingent commissions.

This steering is potentially problematic because the insurers that pay contingent commissions to an intermediary may not best suit a particular insured's risk preferences and needs. At the very least, these insurers will tend to charge more than other insurers in order to support the costs of paying contingent commissions. A recent study found that insurers pass through the entire cost of contingent commissions to their insureds.⁴³ The study controlled for insurers' size, financial leverage, mix of business by line, and diversification across lines of business.⁴⁴ In other words, while insurers pay contingent commissions to intermediaries, consumers ultimately bear the cost of these payments in the form of increased premiums.⁴⁵ Insurers that pay contingent commissions may be suboptimal for reasons other than price, however. These insurers may provide coverage that deviates from a customer's preferences with regard to claims handling, financial strength, risk management services, or coverage scope. Although there is no empirical evidence that insurers paying contingent commissions are systematically worse than other insurers along these quality dimensions, there are certain to be instances in which the best insurance choice for the consumer will not align with the insurance option that pays his or her intermediary the most in contingent commissions. To the extent that independent intermediaries prioritize their receipt of contingent commissions over their market-matching role, then customers will often receive insurance of a quality and price that is not best for their needs.

Of course, sales personnel in a variety of other industries likely face similar incentives to promote specific product types. Although lawmakers do not typically worry about these conflicts of interest, contingent commissions are distinctive for two important reasons. First, unlike a car salesman or a store

43. See Cummins & Doherty, *supra* note 2, at 383. Insurers also pass through the full costs of premium commissions to their insureds. *Id.* These estimates were derived using data that exploited the fact that different insurers offer different levels of contingent commissions. *Id.* at 380-83. It is therefore not inconsistent with the argument presented later, see *infra* notes 168-170 and accompanying text, that insurers may adopt a strategy of charging more to insurance consumers who purchase from intermediaries that completely disclaim contingent commissions.

44. Cummins & Doherty, *supra* note 2, at 380-83.

45. As the authors of the study rightly note, the mere fact that insurers who pay contingent commissions also tend to charge more for their insurance does not necessarily mean that these insurers are worse for the consumer. If the increased costs attributable to contingent commission payments create benefits for the consumer in the form of reduced adverse selection, see *infra* Section I.B., and these benefits outweigh the price increase as well as the other potential costs of contingent commissions, then consumers are better off purchasing insurance from these intermediaries, see Cummins & Doherty, *supra* note 2, at 381-83.

clerk, independent intermediaries actively advertise their independence.⁴⁶ The website of the Independent Agents and Brokers of America encourages consumers to purchase insurance through an independent intermediary rather than through a captive agent or direct underwriter because independent agents are not beholden to any one company.⁴⁷ Rather, the website claims that they are “value hunter[s] who look[] after your pocketbook in finding the best combination of price, coverage and service” because “serving you is your independent agent’s most important concern.”⁴⁸ Similarly, prospective customers are given brochures explaining that their independent agent “shop[s] among various companies” to “find the best combination of coverage, price and service—the best value for your insurance dollar.”⁴⁹ These brochures describe the agent as someone “who . . . goes to bat for you when you need help or advice” and who “is your Personal Insurance Adviser—someone you can count on and trust.”⁵⁰ Each of these claims of independence is confirmed by the simple, but important, fact that independent agents have their own storefronts or office buildings, which advertise the independent agent’s company name, not the names of insurers. This professed independence is one of the primary reasons that insureds choose to purchase their insurance through independent intermediaries.⁵¹ By contrast, consumers understand that sales agents often have an incentive to steer their customers to particular (usually more expensive) products, and discount the advice they receive accordingly.

Second, ordinary consumers typically have little capacity to compare insurance products on their own without the benefits of independent advice from intermediaries.⁵² Insurance is a complicated and intangible product with

46. See HUNTER, *supra* note 13, at 3; *infra* notes 138-140 and accompanying text; see also *Money on the Table*, LEADER’S EDGE, Jan.-Feb. 2007, available at http://www.ciab.com/Content/ContentGroups/Leaders_Edge_Magazine2/2007/Jan_Feb3/Money_on_the_Table.htm (“The distinction between an agent and broker has been lost because most agents present themselves as representing the interests of the insured.” (quoting Bobby Reagan, Reagan Consulting)).

47. See Indep. Ins. Agents & Brokers of Am., Inc., <http://www.iaaa.org/> (follow “Consumer Information” hyperlink; then follow “Choosing an Independent Agent” hyperlink) (last visited April 11, 2007).

48. See *id.*; see also Hunter, *supra* note 11, at 3. Such language is also common in insurance brokers’ materials. Indeed, Marsh & McClellan’s website proclaimed that “[o]ur guiding principle is to consider our client’s best interest in all placements” and added: “We are our clients’ advocates, and we represent them in negotiations. We don’t represent the [insurance companies].” See Marsh Complaint, *supra* note 10, ¶ 6.

49. NAT’L ASS’N OF PROF’L INS. AGENTS, STRAIGHT TALK ABOUT CHOOSING A PROFESSIONAL INSURANCE AGENT (1997), available at <http://www.pianet.com/Publications/choosinganagentbrochure.htm>.

50. *Id.*

51. See, e.g., The Motley Fool, *Insurance Center: Basics*, <http://www.fool.com/insurancecenter/basics/basics07.htm> (last visited Apr. 11, 2007) (“An independent insurance agent represents a number of insurance companies and can more objectively weigh pluses and minuses across many companies and types of insurance.”).

52. See generally Schwarcz, *supra* note 32, at 1412-22 (discussing the avenues by which insurance consumers learn about different insurance options and the limitations of these sources of information).

Beyond Disclosure: The Case for Banning Contingent Commissions

which most consumers have very little experience or knowledge.⁵³ Moreover, insurers' reputations imperfectly reflect the underlying quality of their products and services.⁵⁴ In markets with such limited consumer information, truly independent intermediaries are crucial for market outcomes to be efficient.⁵⁵ Independent advice is significantly less important in markets where products have relatively well-understood features and accurate reputations.⁵⁶

Defenders of contingent commissions admit that the practice theoretically creates conflicts of interest for independent intermediaries, but they argue that competition will minimize the negative impact of these conflicts, especially when supplemented with disclosure. For most independent producers, contingent commissions constitute only a small fraction of their overall revenue, with estimates ranging between 4% and 6% on average.⁵⁷ The vast majority of most intermediaries' compensation instead derives from ordinary premium commissions.⁵⁸ As such, defenders of contingent commissions reason that the primary goal for independent intermediaries is to sell insurance and generate premium-based commissions. Doing so, of course, critically depends on attracting and retaining customers. According to these commentators, were brokers and independent agents systematically to steer their customers to overpriced, low-quality coverage, their loss in premium revenues would far exceed

53. See Robert Keeton, *Insurance Law Rights at Variance with Policy Provisions*, 83 HARV. L. REV. 961, 968 (1970); Mark C. Rahdert, *Reasonable Expectations Reconsidered*, 18 CONN. L. REV. 323, 339 (1986) ("[I]nsurance is an intangible product that is not particularly well understood by the average consumer.").

54. See Schwarcz, *supra* note 32, at 1413-15 (arguing that insurers' reputations do not perfectly reflect the overall quality of their products because, while reputation is largely a function of friends' and families' experience, few insurance consumers receive the most important features of their insurance—protection against low-probability, high-cost losses).

55. See ROBERT H. JERRY II, UNDERSTANDING INSURANCE LAW § 32[b] (2d ed. 1996) ("In forming a contract, an insured relies not upon the text of the policies but on the general descriptions of the coverage provided by the insurer and its agents during the time the insured is considering submitting an application."); Biglaiser, *supra* note 22, at 221 (describing how intermediaries can help facilitate efficient purchasers in markets where information is imperfect); W. David Slawson, *Standard Form Contracts and Democratic Control of Lawmaking Power*, 84 HARV. L. REV. 529, 547 (1971) (noting that the average consumer "depends on an insurance agent and insurance company to sell him a policy that 'works' for its intended purpose in much the same way that he depends on a television salesman and television manufacturer").

56. See Biglaiser, *supra* note 22, at 221.

57. One survey found that, on average, contingent commissions account for only 6% of the revenues received by independent agents. Anne Gron, *Compensation and Industry Profitability: Evidence from the Property-Casualty Insurance Industry*, 71 J. BUS. 407, 410 (1998) (citing INDEP. INS. AGENTS OF AM., AN INDEPENDENT AGENCY'S GUIDE TO PROFIT-SHARING/INCENTIVE COMPENSATION AGREEMENTS (1988)). Other surveys have found similar, though slightly lower, figures for insurance brokers. Cummins & Doherty, *supra* note 2, at 375 n.16 (citing CONNING & CO., COMMERCIAL INSURANCE BROKERS: THEY SNOOZE, THEY LOSE (1999) (finding in a survey of brokers that contingent commissions accounted on average for 5% of brokers' revenues in 1994 and 4.6% in 1999)). For both brokers and independent agents, the percentage of revenue attributable to contingent commissions varies widely. One survey, for instance, found that some brokers received almost 12% of their revenue from contingent commissions, while others received only 1% of their revenue from such commissions. See *id.* at 365 tbl.2.

58. See Cummins & Doherty, *supra* note 2, at 375; Fitzpatrick, *supra* note 3, at 3056.

whatever minimal gain they could receive in contingent commissions.⁵⁹ This is especially true given the competitiveness of most property-casualty insurance markets.⁶⁰ Disclosure of contingent commissions, while not strictly necessary for this competitive process, can help facilitate it. Such disclosure encourages insurance consumers to monitor the quality of their intermediaries' advice or, alternatively, to insist on intermediaries who do not accept contingent commissions.

It is for similar reasons, these commentators have suggested, that few officials or regulators worry about whether ordinary, premium-based commissions will cause intermediaries to steer their customers to particularly expensive insurance. Although intermediaries could theoretically increase their premium-based commissions through steering in the short run, market forces deter this result in the long run.⁶¹ The same logic applies, according to these writers, to contingent commissions. To the extent that insurance costs more when the insurer pays contingent commissions, the best explanation is that this

59. See Cummins & Doherty, *supra* note 2, at 385 (arguing against the notion that intermediaries receiving contingent commissions will steer customers because "contingent commissions account for only about 5 percent of revenue," and "[i]ntermediaries who make inferior placements in pursuit of higher contingent commissions are balancing a small gain against the possibility of a much larger loss, i.e., the loss of the premium-based commission if the client becomes dissatisfied and switches to a competitor"); Fitzpatrick, *supra* note 3, at 3061-62 (noting that concern that intermediaries will steer consumers to suboptimal insurers is not "utterly unfounded," but dismissing it because "insurance intermediaries are normally more concerned with the risk of losing a good customer to a competing producer than they are with any marginal inducements that may be provided by any one insurance carrier").

60. For a response to this argument, see *infra* Section II.A, which argues that competitive forces are unlikely to negate the costs of contingent commissions. This does not mean that the industry as a whole is not competitive on price. See J. DAVID CUMMINS, DEREGULATING PROPERTY-CASUALTY INSURANCE 2-3 (2002) (explaining that deregulation of property-casualty insurance could benefit consumers because "the insurance industry is competitive"); Paul Joskow, *Cartels, Competition and Regulation in the Property-Liability Insurance Industry*, 4 BELL J. ECON. & MGMT. SCI. 375, 391 ("The property-liability insurance industry possesses the structural characteristics normally associated with the idealized competitive market: a large number of firms, operating in a market with low concentration levels, selling essentially identical products, provided at constant unit costs and with ease of entry of new and potential competitors."). For an argument that the industry is not competitive with regard to policy drafting, see Schwarcz, *supra* note 32, at 1401-26.

61. There are at least three important distinctions between the conflicts of interest generated by premium-based commissions and those generated by contingent commissions. First, premium-based commissions create a very simple type of conflict: they mean that producers earn more when their customers pay more. By contrast, because contingent commissions are linked to insurers' profits, they create a complicated array of conflicts that can impact advice about insurance *quality* as well as price. Insurers' profits can increase either if the price of the insurance is too high or if the cost of the insurance is too low. Market forces will generally be better at limiting large price differentials in insurance than at addressing biased advice about quality or the necessity of different insurance options. See *infra* text accompanying notes 114-128. Second, premium-based commissions are much easier to disclose effectively to insurance consumers than are contingent commissions because the former operate only on individual transactions. In contrast to contingent commissions, whose impact on an independent intermediary can only be assessed when looking at aggregate, year-end data, individuals can immediately deduce the relevance of differential premium-based commissions. See *infra* note 119 (explaining that this difficulty in disclosing contingent commissions potentially can be overcome by the use of retrospective supplemental compensation arrangements). Third, premium-based commissions are generally standard across the industry within a given insurance line. See *infra* note 164.

Beyond Disclosure: The Case for Banning Contingent Commissions

extra cost is warranted by the benefits that contingent commissions create for the consumer.⁶²

B. *Contingent Commissions and Adverse Selection*

The core defense of contingent commissions rests on the observation that these payments align the interests of independent intermediaries and insurers: both are better off when the insurer makes a profit on the intermediary's book of business. As two insurance economists, J. David Cummins and Neil A. Doherty, have recently argued in an influential paper,⁶³ this commonality of interests potentially alters intermediaries' behavior in a significant way: it encourages intermediaries to help insurers charge their customers an actuarially appropriate premium. If insurers charge an intermediary's customer a premium that is too low, then the intermediary's loss ratio for that insurer will suffer, reducing the expected amount of the intermediary's contingent commission. If, on the other hand, the insurer misjudges the customer's risk level in the other direction and charges a premium that is too high, the customer may purchase insurance through a different insurer or decide to forego purchasing insurance altogether. The expected value of the intermediary's contingent commissions will decrease in this situation as well.

According to Cummins and Doherty, independent intermediaries that receive contingent commissions can theoretically help insurers set appropriate premiums by informing them about a potential insured's risk profile.⁶⁴ Independent brokers and agents interact directly with potential insureds and often have longstanding relationships with these clients. These intermediaries may therefore possess information about potential insureds that would

62. See Cummins & Doherty, *supra* note 2, at 394.

63. The study was funded by the American Insurance Association. See *id.* at 359 (acknowledging funding from the Association). On the influence of this argument, see, for example, Fitzpatrick, *supra* note 3, at 3060-61; and Andrea Ortega-Wells, *Wharton Study Finds Agents, Brokers Play Critical Role in Buying Process*, *INS. J.*, June 8, 2005, available at <http://www.insurancejournal.com/news/national/2005/06/08/55791.htm>.

64. Cummins & Doherty, *supra* note 2, at 386-89; see also Laureen Regan & Sharon Tennyson, *Agent Discretion and the Choice of Insurance Marketing System*, 39 *J.L. & ECON.* 637, 639 (1996) ("The agent is the first contact the insurer has with a potential policyholder and may be able to obtain information about the customer which would be difficult or costly for the firm to verify."). There is significant evidence that "captive" insurance agents, who work only for one insurer and therefore do not receive contingent commissions, do indeed perform this information-gathering role. For instance, one training manual for an insurer's captive agents instructs them to treat all potential insureds as "suspects" and to "[b]e proactive and selective in your prospecting activity" and "ruthless in getting rid of" bad risks. RICHARD V. ERICKSON ET AL., *INSURANCE AS GOVERNANCE* 226 (2003). Another insurance company manual insists that agents consider "the *desirability* of the client's entire account," and another reminds agents that it is through their "application of ability, knowledge, experience and courage that risks can be selected properly and produce a profit for the insurer." *Id.* at 228. A major work in the sociology of insurance thus finds that agents are "instructed that the selection and rating process is shot through with discretion." *Id.* at 239.

otherwise be unavailable to insurance companies.⁶⁵ In some cases, this information may consist of concrete facts about the insured that are not captured in the ordinary insurance application process. For instance, an intermediary may know that a potential insured was nearly sued earlier in the year, but that the potential plaintiff unexpectedly dropped the suit for personal reasons. Perhaps more frequently, though, an intermediary's informational advantage may consist of subjective impressions about an insurance consumer's risk level that are impossible for the insurer to discern through application forms.⁶⁶ The intermediary may observe, for instance, that a small business owner tends to miss scheduled appointments or looks disheveled and disorganized. Without contingent commissions, Cummins and Doherty suggest, there would be "no incentive for the agent to reveal the information about policyholder risk types" to the insurer unless the insurer engaged in costly monitoring of its agents.⁶⁷

Contingent commissions may improve insurer information even if intermediaries and insurers do not explicitly communicate with one another. As described above, contingent commissions may cause intermediaries to steer relatively high-risk insureds to insurers that do not offer contingent commissions, so as to maintain a good "loss ratio" with those insurers that do.⁶⁸ The better an intermediary's loss ratio is, the more he or she will receive in contingent commissions.⁶⁹ Consequently, contingent commissions may allow insurers to know that they are receiving relatively low-risk insureds from their intermediaries irrespective of their direct communications with these intermediaries.⁷⁰ In this way, contingent commissions may actually create efficient, rather than problematic, steering.⁷¹

65. See Cummins & Doherty, *supra* note 2, at 386.

66. See Regan & Tennyson, *supra* note 64, at 639 ("It is widely acknowledged that agents often employ subjective criteria in evaluating insurance applicants."). Agents selling homeowners insurance, for instance, have been told to identify whether potential insureds live near a "'hangout' for a local youth gang," whether they keep their home clean, or whether a household pet appeared "unusual or vicious." ERICKSON ET AL., *supra* note 64, at 230, 232. As the manual explained, "[p]oorly kept or poorly maintained premises may indicate a lack of responsibility on the potential client's part," and "[o]bvious damage to furniture, carpets, and other personal property can also indicate a moral or liability hazard." *Id.* at 230. One insurance agent operator explained that sometimes he rejected an insurance applicant after meeting with him because he "had a 'gut feeling' that the person was simply 'bad luck.'" *Id.* at 247. Insurers often grant a similar level of discretion to their agents when it comes to automobile insurance, with one auto insurance manual explaining that "[a]gents working in the field are in the best position to know or observe any undesirable characteristics of the clients as operators of the vehicle." *Id.* at 264.

67. Cummins & Doherty, *supra* note 2, at 389. It is modest support for this argument that some insurers offer differential commissions to their captive agents based on the loss ratio of that agent's book of business. See ERICKSON ET AL., *supra* note 64, at 246. The explicit purpose of this scheme is to improve the agent's selections of risks. See *id.* at 246-47.

68. See *supra* note 42 and accompanying text.

69. See *supra* note 37 and accompanying text.

70. For a critical response to this argument, see *infra* notes 157-161 and accompanying text.

71. See *supra* note 42.

Beyond Disclosure: The Case for Banning Contingent Commissions

To the extent that contingent commissions do indeed improve the information that intermediaries transmit to insurers, the increase in efficiency is potentially significant. Decreasing information asymmetries between insurers and insureds is particularly valuable in insurance markets because of the possibility that such asymmetries result in adverse selection.⁷² Adverse selection “is commonly described as the tendency of persons with relatively greater exposure to risk to seek [more] insurance protection.”⁷³ It can occur when insurance consumers have information about their potential risk that insurers cannot observe, making it difficult for insurers to offer insurance to low-risk individuals at an actuarially fair price. This potentially results in a self-reinforcing trend: low-risk insureds tend to forego insurance because it is too expensive for them, resulting in mostly high-risk insureds purchasing insurance, resulting in a further price increase.⁷⁴

C. Other Potential Benefits of Contingent Commissions

Any productive assessment of contingent commissions must compare the risk that the practice biases intermediaries’ advice against the prospect that it improves insurer information and thus limits adverse selection.⁷⁵ To date, most of the other arguments concerning contingent commissions that commentators

72. See George Priest, *The Current Insurance Crisis and Modern Tort Law*, 96 YALE L.J. 1521, 1541 (1987) (“Adverse selection is a problem central to every insurance context, and it dominates the insurance function.”). The seminal article describing adverse selection in the insurance market is Michael Rothschild & Joseph Stiglitz, *Equilibrium in Competitive Insurance Markets*, 90 Q.J. ECON. 629 (1976).

73. Priest, *supra* note 72, at 1541. Priest explains the adverse selection problem as follows:

An insurer must collect into a risk pool individuals with a sufficiently narrow range of exposure to risk for the insurance to remain financially attractive to each member of the pool. The insurance premium for the pool must be set according to the average level of risk brought to the pool. The wider the range between high-risk and low-risk pool members, the greater the difference between the average risk and the risk of low-risk members. Low-risk members pay a premium that, because it is based on an average which includes high-risk members, is more than they would have to pay if they could be segregated into a risk pool of their own. If the disparity between the premium and the risks added by low-risk members becomes too substantial, low-risk members may drop out of the pool because they find alternative means of protection cheaper than market insurance.

Id.

74. A secondary benefit of facilitating credible communication between independent intermediaries and insurers, which Cummins and Doherty do not discuss, is that doing so may improve the efficiency of the insurance investigation process. Insurance intermediaries can more efficiently inquire about a client’s risks than an insurer because intermediaries need only do so once in order to transmit those conclusions to multiple different insurers. Regan & Tennyson, *supra* note 64, at 639. By contrast, if the entire process of vetting the risks associated with potential insureds were left to each insurer, then multiple insurers would need to conduct repetitive inquiries whenever an insured was turned down for coverage by the first insurer. If contingent commissions facilitate intermediaries’ willingness and ability to transmit accurate information to insurers, then it may shift parts of the investigative process from insurers to intermediaries and thus reduce the number of duplicative inquiries that need to be made. *See id.*

75. One additional relevant consideration, introduced *infra* Section III.B, is that contingent commissions may induce strategic, socially inefficient behavior by high-risk, sophisticated insureds.

and advocates have advanced are largely unconvincing and serve to muddy the key analytical issues. This Section reviews these issues and explains why they ought not to factor significantly in the ultimate policy analysis.

1. *Contingent Commissions as a Method of Expanding Insurance Coverage*

In a recent paper, two leading insurance economists argued that contingent commissions may be desirable because they can help expand the insurance market to cover theoretically insurable losses that are not identified in the governing policy.⁷⁶ In a number of circumstances, insurers can efficiently cover insureds' losses even though the applicable policy does not provide such coverage.⁷⁷ Despite the exclusion of these losses from policy terms, the paper suggests that independent intermediaries may be able to induce insurers to pay such nonverifiable losses ex post when doing so is efficient: not only do independent intermediaries have the expertise to decide when insurance can efficiently be provided for nonverifiable losses, but they have the power to remove their customers from an insurer that fails to pay for such losses.⁷⁸ But intermediaries, according to the argument, will only have a strong incentive to police coverage for nonverifiable losses to the extent that they receive some slice of the value that is generated by extending insurance in this way.⁷⁹ Contingent commissions, the article concludes, can serve this role because the intermediary receives commissions in proportion to the insurer's profitability, which includes the additional rent the insurer charges for covering nonverifiable losses.⁸⁰

This argument is unpersuasive for two reasons. First, it is unlikely that contingent commissions actually enhance independent intermediaries' willingness to punish insurers who do not provide coverage for nonverifiable

76. Doherty & Muermann, *supra* note 23.

77. *See id.* at 3-4; Schwarcz, *supra* note 32, at 1401-26 (arguing that insurance policies may fail to provide efficient coverage due to various market failures in the drafting of insurance policies). Such events may not have been anticipated, may have been too complicated to include in the contract, or may have been excluded from the contract due to market imperfections. *See* Doherty & Muermann, *supra* note 23, at 3-4.

78. *See* Doherty & Muermann, *supra* note 23, at 6-8. Interestingly, this theory is consistent with a recent article in which Jason Scott Johnston has argued that, in a variety of contexts, standard form contracts facilitate ex post bargaining between consumers and company employees who have the authority to grant exceptions from the terms of these contracts. *See* Jason Scott Johnston, *The Return of Bargain: An Economic Theory of How Standard-Form Contracts Enable Cooperative Negotiation Between Businesses and Consumers*, 104 MICH. L. REV. 857 (2006). Of course, Doherty and Muermann's claim is not simply that insurers grant exceptions from insurance policy terms ex post, but that independent insurance agents and brokers facilitate that role. It is this latter claim that is rejected above.

79. Doherty & Muermann, *supra* note 23, at 9 ("In our model, brokers seek future efficiency gains because they can capture rents directly from value added.")

80. *See id.* at 8-9.

Beyond Disclosure: The Case for Banning Contingent Commissions

losses ex post. The fact that intermediaries receive contingent commissions actually creates an incentive for them *not* to punish insurers who do not cover nonverifiable losses ex post: insurers profit more from this result, meaning that the contingent commissions that intermediaries receive are larger.⁸¹ Admittedly, contingent commissions may encourage intermediaries to claim, ex ante, that they will police nonverifiable losses. But they—just like insurers—will invariably face the opposite incentive ex post. In the end, the only reason that the article concludes otherwise is that it assumes that insureds are fully aware of intermediaries' reputations for negotiating coverage for nonverifiable losses ex post.⁸² Yet this assumption about insurance intermediaries is no more plausible than the assumption that insureds can directly observe insurance companies' reputations for ex post settlement of nonverifiable claims. The authors specifically reject this "informational assumption" as too "strong."⁸³ In truth, though, insureds are likely to be *better* able to observe insurers' tendencies to pay nonverifiable claims ex post than to observe intermediaries' tendency to advocate for this result.⁸⁴ The latter cannot easily be inferred from claims histories, and significantly more information is available about insurers' reputations than intermediaries' reputations: there are fewer insurers than intermediaries, and information about insurers' payment practices and service levels are readily available from consumer magazines and websites.⁸⁵

A second limitation of this argument is that it is unclear why contingent commissions are any better at inducing intermediaries to facilitate the coverage of nonverifiable losses than ordinary premium-based commissions. If insurance premiums are higher because insurance coverage has been expanded to include coverage for nonverifiable losses, then intermediaries will be compensated through higher premium-based commissions. Contingent commissions, then, are not necessary for the underlying argument about nonverifiable coverage to work.

81. See *supra* note 41.

82. See Doherty & Muermann, *supra* note 23, at 14 (stating the assumption that "[e]ach policyholder observes [the broker's] settlement of a non-verifiable claim, or the broker's penalty on the insurer if it is not paid").

83. *Id.* at 13 (arguing that while brokers are not "strictly necessary to create a market for non-verifiable losses," the "informational assumptions are strong" because "all policyholders must observe all of the incumbent insurers [sic] claim payments to all policyholders and be prepared to switch if even one policyholder is denied").

84. The fact that insurance purchasers can evaluate the reputations of insurers better than the reputations of intermediaries does not mean that, in absolute terms, insurance consumers are generally informed about insurers' reputations. See *supra* note 54 and accompanying text.

85. See Schwarcz, *supra* note 32, at 1419 (describing how insurance consumers rely on secondary material to learn about their insurance options); see also *supra* note 29 (citing *Consumer Reports* articles discussing insurers' reputations for claims handling).

2. *Contingent Commissions as a Protector of Small Agencies*

Another argument advanced in favor of contingent commissions is that they are necessary to prevent the “bankrupt[cy] [of] hundreds of small insurance agencies in communities throughout America” and the consequent “further consolidation of insurance brokerage business in large global firms.”⁸⁶ The logic of this argument is that contingent commissions function as “house money” that is necessary for insurance agencies to pay for their basic overhead expenses.⁸⁷ According to this argument, without contingent commissions, a disproportionate amount of money would be retained by employee-agents of small agencies. The reason is that employee-agents have disproportionate bargaining power over the agencies that employ them: employee-agents can easily move to competing agencies because they own their “book of business,” allowing them to take their existing clients with them to another agency if they move themselves.⁸⁸ Contingent commissions ensure the viability of small insurance agencies in the face of this asymmetric bargaining power, the argument goes, because they are typically paid only at year’s end and are thus largely invisible to the agency’s employee-agents.⁸⁹ As such, these commissions can be surreptitiously retained by the agency and used to pay for its basic expenses.

The argument that banning contingent commissions will bankrupt an inefficiently large number of small independent insurance agencies ignores economic forces in two related ways. First, it is highly unlikely that insurance agents’ compensation depends on the visibility to agents of their employers’ profits. In competitive labor markets,⁹⁰ employees’ wages are set by their marginal value to their firms. Employees need not know precisely how much they contribute to their employers’ bottom line for this to be so: if they contribute more than they are paid, then a competing firm will offer them a higher wage. If

86. Fitzpatrick, *supra* note 3, at 3042; see also Albert Lloyd & Craig Niess, *The Impact of Contingent Commissions on Independent Agents*, INS. J., Mar. 12, 2007, available at <http://www.insurancejournal.com/magazines/east/2007/03/12/features/78016.htm> (“The very ability of an agency to service and support itself hinges on the fact that contingents are considered a legitimate revenue stream.”). Fitzpatrick does not argue that contingent commissions are necessary for the survival of small, niche insurance brokers. See Fitzpatrick, *supra* note 3, at 3059. The reason, presumably, is that the brokerage market is much more consolidated than the insurance agency market, and so it seems less plausible that independent agents can demand salaries in excess of their value to their employer-brokers.

87. See Fitzpatrick, *supra* note 3, at 3058-59.

88. *Id.* at 3058.

89. See *id.* at 3059 (arguing that smaller, independent intermediaries have long attempted to “preserve the confidentiality of contingent commission arrangements” in order to “keep[] such information from the agency’s own employees,” because “contingent commissions provide a vital pool of ‘house’ money that funds the basic overhead of their firms,” without which many of these intermediaries could not remain in business).

90. The argument discussed in this Subsection clearly rests on the assumption that markets for insurance intermediaries are indeed competitive. To the extent they are not, then the underlying argument that contingent commissions are necessary to protect small insurance agencies falls flat.

Beyond Disclosure: The Case for Banning Contingent Commissions

they contribute less, they will be fired or their wage will be decreased.⁹¹ Second, if banning contingent commissions would indeed bankrupt some small insurance agencies for the reasons described above, then this would probably be a good outcome. The premise underlying this defense of contingent commissions is that some insurance agencies have only been able to maintain their presence in the marketplace by concealing from their employees the true extent to which these employees contribute to overall profits. But if this is so, then the costs incurred by agencies in supporting their agents—by providing advertising, coordinated filing services, office space in which to meet customers, and the like—must be less than the value of those expenditures. And in that case, the efficient outcome would be for the individual agent to move to another agency that adds enough value to the overall transaction to cover its costs.⁹²

3. *Contingent Commissions as a Protector of Small Insurers*

A third defense of contingent commissions is that they may allow relatively small insurers to enter the market more easily by inducing independent intermediaries to steer customers to new insurers. Although intermediaries might ordinarily stay away from new insurers that have not developed a reputation, contingent commissions can induce them to go with these insurers by offering them what amounts to a “bonus” payment. Ultimately, the argument goes, this is beneficial for the competitiveness of the insurance market because it reduces barriers to entry.⁹³

This argument in favor of contingent commissions is also limited, though its recognition of the impact that contingent commissions can have on intermediaries’ behavior is noteworthy. In many markets where reputation is important, new entrants may find it advantageous to provide heavy discounts that cannot be sustained in the long run. This strategy allows new firms to gain some initial market penetration and, more importantly, to credibly signal to consumers that the firm believes in its products: after all, the firm’s strategy is a losing one if consumers who initially move to the firm to take advantage of its low prices do not then stay with the firm once prices have normalized.⁹⁴ But there is little reason why such promotional discounts need be made to insurance intermediaries, as opposed to insureds themselves in the form of lower prices. To the extent that new insurers offer lower premiums that offset their weaker

91. See WILLIAM J. BAUMOL & ALAN S. BLINDER, *ECONOMICS: PRINCIPLES AND POLICY* 378 (10th ed. 2005).

92. The only potential reason this might not be true is if the resulting industry structure were uncompetitive. But it is hard to imagine that, even if contingent commissions do artificially prop up some insurance agencies, banning them would fundamentally alter the composition of the marketplace such that a competitive marketplace would be transformed into an uncompetitive marketplace.

93. Cummins & Doherty, *supra* note 2, at 385-86.

94. See William J. Baumol, *Predation and the Logic of the Average Variable Cost Test*, 39 J.L. & ECON. 49, 68 (1996).

reputation, at least some intermediaries are likely to recommend this option to their customers. And while this strategy may not be as effective as contingent commissions in quickly recruiting new clients, such reduced premiums signal equally well to the marketplace that the insurer believes in the product that it provides.⁹⁵ Because this signal, rather than the initial influx of business that accompanies it, is the crucial mechanism by which new entrants establish themselves with low initial rates, contingent commissions are not necessary to reduce barriers to entry in insurance markets.

4. *Contingent Commissions as a Facilitator of Limited Insurer-Intermediary Interactions*

A final argument in favor of contingent commissions is that these payments may promote efficiency by inducing agents to focus on selling the policies of a limited number of insurers. In many ways, this argument is simply the flip side of the steering argument described above.⁹⁶ Just as contingent commissions may cause intermediaries to steer their customers to suboptimal insurance, they may also cause intermediaries to focus their sales efforts on a few insurers who pay the largest contingent commissions. This may reduce transaction costs: it is presumably cheaper for agents to focus on the mechanics of selling only a few insurers' products. Insurers, as well, may benefit from receiving customers through a limited array of familiar independent intermediaries, assuming the number of customers remains constant.

This defense of contingent commissions is ultimately insignificant. Independent intermediaries certainly incur a cost when they choose to sell an additional insurer's policy line. But contingent commissions are not necessary to ensure that intermediaries appropriately account for this cost in their decision-making. Intermediaries directly bear this cost, whether or not they are paid contingent commissions. By contrast, if there are costs to the insurer from selling policies to the customers of an unfamiliar intermediary, then it is true that the intermediary may not take these costs into account in deciding how to direct its customers. If so, then contingent commissions might promote efficiency. Yet it is hard to imagine that the non-underwriting costs⁹⁷ to an insurer of receiving a client from a relatively less familiar intermediary are substantial. By necessity, insurers that sell their products through independent intermediaries must be adept at the mechanics of providing coverage for

95. Even if intermediaries could not be trusted to recommend insurers who provided promotional discounts on premiums, there is no reason why new insurers could not compensate intermediaries using methods other than contingent commissions, such as higher premium-based commissions.

96. See *supra* Section I.A.

97. To the extent that insurers who distribute their policies through independent intermediaries do enjoy underwriting efficiencies from dealing with familiar intermediaries, this argument collapses into the adverse selection argument, described in Section I.B, *supra*.

Beyond Disclosure: The Case for Banning Contingent Commissions

customers referred by hundreds of different intermediaries. This is true whether or not the insurer offers contingent commissions: even an insurer that pays contingent commissions to only a few independent agents will receive business from hundreds of independent intermediaries.⁹⁸

II. CONTINGENT COMMISSIONS IN CONSUMER INSURANCE MARKETS

The emerging consensus among lawmakers and commentators is that, given the competing costs and benefits of contingent commissions, these payments should be permitted in consumer insurance markets so long as they are appropriately disclosed to consumers.⁹⁹ Such a disclosure regime is primarily intended to limit the core potential cost of contingent commissions—that they cause intermediaries to steer consumers to suboptimal insurance.¹⁰⁰ At the same time, a disclosure-based response does not interfere with the potential of contingent commissions to reduce adverse selection.

This Part challenges the argument that disclosure is a sufficient regulatory or legislative response to contingent commission payments in consumer insurance markets. Drawing on the experiences of regulators in addressing other trilateral dilemmas, Section A argues that the core cost of contingent commissions—the risk of steering—is both significant and largely intractable when consumers purchase insurance through intermediaries that accept

98. Although the marginal cost of processing customers from an entirely new intermediary may be significant, this is likely to be a one-time cost for the insurer: the intermediary's name and address must be placed in the computer, and the intermediary will likely have to sign a distribution contract. However, the effect that contingent commissions have on the ultimate number of independent intermediaries from which an insurer receives business is likely trivial.

99. The leading model legislation, drafted by the National Conference of Insurance Legislatures, requires disclosure only when a producer is paid by both the insurance consumer and the insurer—a rare occurrence. That disclosure must acknowledge that “compensation [from an insurer] will be received by the producer or affiliate,” and the customer must receive “a description of the method and factors utilized for calculating the compensation to be received from the insurer or other third party for that placement.” PRODUCER COMPENSATION DISCLOSURE MODEL, AMENDMENT TO THE PRODUCER LICENSING MODEL ACT § 1(A)(1) (Nat'l Conference of Ins. Legislators 2005), <http://www.aba.com/aba/documents/abia/NCOILFinalMarkup.pdf>. Another, more effective proposal would inform the insurance consumer in all circumstances that:

OUR FIRM HAS AGENCY CONTRACTS WITH THE FOLLOWING INSURANCE COMPANIES: _____

IF YOU CHOOSE, OUR FIRM WILL ACT AS A “DUAL AGENT,” REPRESENTING BOTH YOU AND THE INSURERS WHO HAVE APPOINTED US, IN YOUR INSURANCE PURCHASE. IF WE ACT AS A “DUAL AGENT,” WE WILL BE COMPENSATED ON A COMMISSION BASIS, WITH OUR COMMISSION BEING INCLUDED IN YOUR POLICY PREMIUM AND PAID BY THE INSURER YOU SELECT. WE MAY ALSO BE ELIGIBLE TO RECEIVE ADDITIONAL COMPENSATION FROM THAT INSURER BASED ON THE OVERALL VOLUME AND PROFITABILITY OF THE POLICIES WE WRITE WITH THAT INSURER. INFORMATION ABOUT SUCH ADDITIONAL COMPENSATION FOR WHICH OUR FIRM MAY BE ELIGIBLE CAN BE FOUND ON OUR WEB SITE, AND IS ALSO PROVIDED BY THE INSURERS WE REPRESENT ON THEIR WEB SITES.

Fitzpatrick, *supra* note 3, at 3069-70.

100. See *supra* Section I.A.

contingent commissions. The mere disclosure of contingent commissions does not meaningfully alter this risk, which is largely a function of inherent informational asymmetries between consumers and intermediaries. Nor does disclosure of any type allow consumers to assess meaningfully the magnitude or probability of this risk.¹⁰¹ Section B continues by arguing that the primary theoretical benefit of contingent commissions—their capacity to mitigate adverse selection—is insignificant in consumer insurance markets. Even if adverse selection were a serious concern in these markets, contingent commissions might not actually improve the information about consumer-insureds that insurers receive from independent intermediaries. Weighing the core costs and benefits of contingent commissions, Section C concludes that regulators or legislatures should ban these payments in consumer insurance markets.

A. *Disclosure, Steering, and Trilateral Dilemmas*

Although contingent commissions have only recently emerged as a significant issue for regulators and legislatures, government officials have frequently addressed analogous consumer market problems, known as “trilateral dilemmas.” In a trilateral dilemma, market intermediaries extract side payments from other professionals in exchange for steering business to them.¹⁰² Trilateral dilemmas involve three parties: a consumer, a market intermediary that influences the consumer’s decisions about hiring additional service providers, and a third-party service provider. In these situations, the third-party service provider may make side payments to the market intermediary in exchange for its referral of business.¹⁰³ For instance, when consumers purchase a home, they typically rely on a number of market intermediaries, such as realtors and mortgage brokers, who themselves provide advice on choosing additional service providers such as lawyers and lending institutions. In such circumstances, the intermediary (the realtor or mortgage broker) may advise the customer to choose a particular service provider (a lawyer or lending institution) based on whether that service provider makes a side payment to the intermediary.¹⁰⁴ Another prominent example of a trilateral dilemma is the potential that investment managers (market intermediaries) will select particular brokerage firms to execute trades on behalf of the investor based on

101. These arguments thus pose a direct challenge to the claim made by defenders of contingent commissions that competition in the insurance industry will eliminate the inefficiencies associated with contingent commissions. *See supra* notes 57-62 and accompanying text.

102. Jackson & Burlingame, *supra* note 25.

103. *See id.*

104. *See id.* One recent and controversial example of this concerns the propriety of mortgage brokers receiving “yield-spread” premiums from their lending institutions. *See also* Kathleen C. Engel & Patricia McCoy, *A Tale of Three Markets: The Law and Economics of Predatory Lending*, 80 TEX. L. REV. 1255, 1264 (2002).

Beyond Disclosure: The Case for Banning Contingent Commissions

the brokerage firm's side payments to the investment manager.¹⁰⁵

In many, if not most, trilateral dilemmas, legislatures and regulators have determined that the risk that intermediaries will provide biased advice to their customers is significant, and that it cannot be sufficiently ameliorated by using disclosure to foster competition.¹⁰⁶ For instance, one of the goals of the Real Estate Settlement Procedures Act (RESPA) of 1974¹⁰⁷ is to protect consumers from exploitation associated with biased advice from real estate intermediaries.¹⁰⁸ To do so, section 8 of RESPA not only erects certain disclosure requirements, but also flatly prohibits a variety of side payments to real estate settlement providers.¹⁰⁹ Similarly, section 28(e) of the Securities Exchange Act of 1934¹¹⁰ is partly intended to limit the risks associated with side payments to investment managers.¹¹¹ It not only requires that investment managers disclose the side payments that they receive from brokerage firms, but also limits the types of side payments that can be made: brokerage firms can only provide side payments to investment managers in the form of "brokerage and research services" that redound to the benefit of clients.¹¹² Recently, public pressure to eliminate even these very narrow side payments has increased.¹¹³ Other instances of trilateral dilemmas that are regulated in ways that extend beyond mere disclosure include employers' choices of 401(k) providers and bankers' arrangements with telemarketers.¹¹⁴

For each of these trilateral dilemmas, there are three basic reasons why disclosure has only a limited capacity to facilitate competition and thus to protect consumers from receiving biased advice. First, for disclosure to allow for an effective market response, consumers must have a basic understanding of

105. See Jackson & Burlingame, *supra* note 25.

106. See *id.*

107. 12 U.S.C. §§ 2601-2617 (2000).

108. See Jackson & Burlingame, *supra* note 25. As Jackson and Burlingame explain, *see id.*, RESPA was partially motivated by a government study that concluded that the payment of kickbacks in real estate sales was "widely employed, rarely inure[s] to the benefit of the home buyer, and generally increase[s] total settlement costs." U.S. DEP'T OF HOUS. & URBAN DEV. & THE VETERANS ADMIN., HUD-F-5, REPORT ON MORTGAGE SETTLEMENT COSTS 5 (1972). Although the study actually recommended that Congress set maximum settlement costs, Congress ultimately took a less drastic measure in RESPA by prohibiting kickbacks and unearned fees and creating a private right of action. See 12 U.S.C. § 2607. For a popular account of the conflicts of interest associated with real estate brokers, see STEVEN D. LEVITT & STEPHEN J. DUBNER, *FREAKONOMICS* 55-89 (2005).

109. 12 U.S.C. § 2607.

110. 15 U.S.C. § 78bb(e) (2000).

111. See Lee B. Burgunder & Karl O. Hartmann, *Soft Dollars and Section 28(e) of the Securities Exchange Act of 1934: A 1985 Perspective*, 24 AM. BUS. L.J. 139 (1986); D. Bruce Johnsen, *Property Rights to Investment Research: The Agency Costs to Soft Dollar Brokerage*, 11 YALE J. ON REG. 75, 83 (1994); Jonathan R. Macey, *Positive Political Theory and Federal Usurpation of the Regulation of Corporate Governance: The Coming Preemption of the Martin Act*, 80 NOTRE DAME L. REV. 951, 971 (2005).

112. 15 U.S.C. § 78bb(e)(1).

113. See Benn Steil, Op-Ed, *The Soft Dollar Scandal*, WALL ST. J., June 10, 2006, at A15.

114. Jackson & Burlingame, *supra* note 25.

the underlying services they are receiving and the prices those services should cost.¹¹⁵ Otherwise, consumers will be unable to assess whether their end service providers are charging too much or providing too little, even if they know that this is potentially the case. In most trilateral dilemmas, however, consumers do not understand much about the underlying services they are purchasing; that is precisely why they seek out the advice of an intermediary. Second, even when disclosure effectively informs some sophisticated consumers, competition cannot protect the interests of other, less sophisticated consumers if intermediaries can distinguish between these two groups and can offer advice accordingly.¹¹⁶ Such discrimination is easy in the case of many trilateral dilemmas, as intermediaries typically know a lot about their customers and are specifically trained to provide individualized advice to them.¹¹⁷ Third, disclosure may be ineffective when intermediaries have a position of power over—or a relationship of trust with—consumers, as they often do in trilateral dilemmas.¹¹⁸ In such circumstances, consumers will be more likely to ignore the prospect that financial incentives will cause an intermediary to offer biased advice.

These explanations for the limited capacity of disclosure and competition to protect consumers in most trilateral dilemmas apply squarely to contingent commissions in consumer insurance markets. First and most importantly, insurance consumers typically have a limited understanding of the underlying insurance transaction that no conceivable disclosure could remedy. *Even assuming that consumers would read and understand a contingent commission*

115. *Id.* at 43; cf. Steven L. Schwarcz, *Rethinking the Disclosure Paradigm in a World of Complexity*, 2004 U. ILL. L. REV. 1 (arguing that certain firms' financial arrangements, such as those of Enron, are so complex that regulation should go beyond a disclosure-based paradigm).

116. See R. Ted Cruz & Jeffrey J. Hinck, *Not My Brother's Keeper: The Inability of the Informed Minority To Correct for Imperfect Information*, 47 HASTINGS L.J. 635, 675 (1998); Avery Wiener Katz, *Standard Form Contracts*, in 3 THE NEW PALGRAVE DICTIONARY OF ECONOMICS AND THE LAW 502, 504-05 (Peter Newman ed., 1998).

117. For instance, litigation has uncovered various cases of such discrimination by lending agents. See ELIZABETH WARREN & AMELIA WARREN TYAGI, *THE TWO INCOME TRAP* 135-36 (2003) (quoting one loan officer who explained that "if someone appeared uneducated, inarticulate, was a minority, or was particularly old or young, I would try to include all the [additional costs] CitiFinancial offered" (alteration in original)). Another well-documented example of this phenomenon occurs in the automobile context, where salesmen are given wide latitude to negotiate prices so that less sophisticated consumers who do not know better can be charged more, while the business of more sophisticated consumers can be retained by significantly lowering the car's price below the sticker amount. See Ian Ayres, *Fair Driving: Gender and Race Discrimination in Retail Car Negotiations*, 104 HARV. L. REV. 817, 845 (1991) (finding evidence that automobile retailers sell cars for higher prices to minorities, and suggesting that part of the reason may be that sellers use race as a proxy for how much consumers shop for better prices); Ian Ayres, *Further Evidence of Discrimination in New Car Negotiations and Estimates of Its Cause*, 94 MICH. L. REV. 109, 138-40 (1995) (similar).

118. See Jackson & Burlingame, *supra* note 25. Jackson and Burlingame note an additional reason why market forces may not correct for side payments in many trilateral dilemmas: intermediaries typically purchase a wide array of settlement services, and the consumer has a limited capacity to monitor all of these at once. This is less of an issue in the insurance context, as the intermediary generally only arranges for a single service provider. See *id.*

Beyond Disclosure: The Case for Banning Contingent Commissions

disclosure—a generous and limited assumption¹¹⁹—they will be ill-equipped to police the underlying conflict of interest or to assess its significance. Insurance consumers generally do not know how to assess the quality of different insurance options: the risks associated with an insurer's "fair" financial rating, for instance, are beyond the ken of most insurance consumers, including many small businesses and otherwise savvy individuals.¹²⁰ Similarly, most insurance consumers have no idea which insurers have a poor reputation for claims management or use insurance policies that are narrower than the standardized coverage forms.¹²¹ Indeed, it is precisely for these reasons that consumers choose to rely on independent insurance intermediaries.¹²²

Without the ability to assess the quality of different insurance options, even consumers who are aware of contingent commissions are susceptible to being steered to insurers that are too expensive, of low quality, or otherwise inappropriate. Although informed consumers can compare the prices available through different intermediaries or the relative commissions that intermediaries receive, these comparisons are largely unhelpful when they are decoupled from the ability to assess the quality of the associated insurance options. Insurance may appear reasonably priced even if it is actually quite expensive for the level of coverage and services provided, while more expensive insurance may in fact be the best deal for the consumer.¹²³ Similarly, the mere fact that an intermediary receives a large contingent commission from a recommended insurer hardly means that her advice is necessarily suspect. Without some

119. See Robert S. Adler & R. David Pittle, *Cajolery or Command: Are Education Campaigns an Adequate Substitute for Regulation?*, 1 YALE J. ON REG. 159, 161 (1984); W. Kip Viscusi, *Individual Rationality, Hazard Warnings, and the Foundations of Tort Law*, 48 RUTGERS L. REV. 625, 650 (1996). In order for disclosures to be effective, consumers must, of course, read and understand them. But consumers often do not read long and complicated disclosures, and frequently do not understand them when they do. See Howard Latin, "Good" Warnings, Bad Products, and Cognitive Limitations, 41 UCLA L. REV. 1193, 1242 (1994). To the extent that lawmakers do indeed opt for a disclosure-based solution, supplemental compensation arrangements can significantly improve the effectiveness of disclosure by reducing the complexities of contingent commission arrangements to a single number that indicates the size of the intermediaries' conflict of interest. See *infra* notes 163-164 and accompanying text.

120. Cf. James M. Fischer, *Why Are Insurance Contracts Subject to Special Rules of Interpretation?: Text Versus Context*, 24 ARIZ. ST. L.J. 995, 1055 (1992) (noting that insureds are generally uninformed about their insurance needs).

121. See Schwarcz, *supra* note 32, at 1413-15 (describing the limits of reputation as a vehicle for informing insurance consumers about the quality of their insurance options); *id.* at 1417-18 (describing how some mega-insurers can use policy language that is narrower than industry norms because they can generate their own actuarial data and enjoy other economies of scale in drafting policy language).

122. See Slawson, *supra* note 55, at 547 ("[The average consumer] depends on an insurance agent and insurance company to sell him a policy that 'works' for its intended purpose in much the same way that he depends on a television salesman and television manufacturer.").

123. This variation in product quality, in addition to price, is part of what distinguishes the conflicts of interests associated with contingent commissions from those associated with premium commissions. See *supra* note 61 and accompanying text. Unlike contingent commissions, premium commissions do not create any incentive for intermediaries to sell insurance of similar price but lower quality as compared to alternative policies.

capacity to independently evaluate the intermediary's advice, price and commission information is of limited use.¹²⁴

It is precisely because insurance is so difficult for most consumers to understand that the dominant method of regulating the industry is through mandatory requirements and prohibitions rather than disclosure-based schemes. Insurance policies are in many ways the most complex financial services that firms offer on the open market because their value is contingent on the insurer's underlying financial stability, the content of the policy itself, and events unrelated to the insurer's assets or the governing contract.¹²⁵ Contrast these characteristics with securities and mutual funds, for example, whose value can be ascertained based solely on the governing contract and the companies' assets.¹²⁶ This differential in inherent complexity explains why securities and mutual funds are typically regulated via relatively nonintrusive measures (such as disclosure and general standards of conduct), whereas insurance regulation often takes the form of specific requirements or prohibitions (such as capital requirements).¹²⁷ Recall, however, that when it comes to trilateral dilemmas in the mutual fund industry, the regulatory strategy of the Securities Exchange Act has uncharacteristically gone beyond disclosure because of concerns that investors could be steered to overly expensive brokerage services.¹²⁸ If the opaqueness of the services that brokerage firms

124. Spitzer's investigation of Marsh & McLennan provides modest evidence of the insufficiency of disclosure when insurance purchasers have a limited understanding of the underlying insurance product. *Cf.* Fitzpatrick, *supra* note 3, at 3047-52 (suggesting that Marsh was unique and that its bid rigging cannot be extrapolated to the rest of the industry). These investigations revealed that Marsh brokers solicited purposefully inflated price quotations from some insurers in order to steer clients to other insurers and to maximize Marsh's receipt of contingent commissions. *See* Marsh Complaint, *supra* note 10, ¶¶ 43-66. Marsh's customers were highly sophisticated firms, at least some of whose risk managers were undoubtedly aware of contingent commission payments. *See* Fitzpatrick, *supra* note 3, at 3049 n.33 (noting that insurers publicly report the payment of contingent commissions in their annual statements filed with the state insurance department). Yet Marsh's bid-rigging rendered these firms unsophisticated as to the appropriate pricing of their policies, enabling Marsh to steer them to suboptimal insurance. The fact that Marsh employees were willing to take this risk may, as others have pointed out, have been an unfortunate result of the company's internal structure. *See* Fitzpatrick, *supra* note 3, at 3047 (noting that Marsh's Global Broking division—the division in which all of the bid-rigging took place—derived its income solely from its receipt of contingent commissions). But the fact that they were able to succeed, at least for a significant period of time, suggests how easy it can be for knowledgeable intermediaries to offer questionable, debatable, or even flatly wrong advice to customers who may be abstractly aware of a potential conflict of interest. *See* Hunter, *supra* note 11, at 2 ("If large sophisticated consumers of insurance can be easily cheated and overcharged, unsophisticated individual and small business buyers are even more vulnerable to such sharp practices.")

125. *See* Howell E. Jackson, *Regulation in a Multisector Financial Services Industry: An Exploratory Essay*, 77 WASH. U. L.Q. 319, 330, 356-57 (1999) (explaining that insurance is a particularly complex financial arrangement because the assets and liabilities of insurance companies are hard to determine).

126. *See id.* at 323-25 (securities); *id.* at 327-28 (mutual funds). *See generally* HOWELL E. JACKSON & EDWARD L. SYMONS, JR., *REGULATION OF FINANCIAL INSTITUTIONS* (1999) (exploring differences in the regulation of various financial services industries and suggesting that many of these differences are attributable to the level of complexity associated with the target industry).

127. *See* Jackson, *supra* note 125, at 353-60.

128. *See supra* notes 111-113 and accompanying text.

Beyond Disclosure: The Case for Banning Contingent Commissions

provide for investors warrants a solution to that trilateral dilemma beyond disclosure, then the same should presumably be true in the context of a more complex financial service.

The second reason that market forces do not protect consumers from most trilateral dilemmas—that intermediaries can distinguish between sophisticated and unsophisticated consumers—also applies in the contingent commission context. When intermediaries can identify unsophisticated consumers, they can steer them to less qualified service providers who make lucrative side payments to the intermediary while directing sophisticated purchasers to objectively superior insurers. Intermediaries can thereby exploit unsophisticated consumers without suffering any serious market penalties for doing so.¹²⁹ Insurance intermediaries are particularly likely to be able to identify and steer uninformed insurance customers in this way because of the close contact they have with their clients. Disclosure does nothing to limit this risk. In fact, disclosure may actually help intermediaries to assess the sophistication and responsiveness of their customers by providing a test of an insurance consumer's relative skepticism and sophistication. Insureds who read a disclosure and ask follow-up questions displaying an understanding of the underlying conflict of interest are relatively less likely to overlook inefficient steering to suboptimal insurers. By contrast, consumers who quickly sign a disclosure without reading the text, or who seem to misunderstand the meaning of the disclosure, will be much more enticing targets for self-interested producers predisposed to steering.

An empirical study of an Arizona independent insurance agency substantiates the risks of such discrimination against unsophisticated consumers by insurance intermediaries.¹³⁰ The agency, which remained unidentified, employed eight agents with no ownership stake in the company and three "equity agents" who received a portion of the agency's profits.¹³¹ Because the contingent commissions the company received were paid directly to the company, the three equity agents stood to gain more from maximizing contingent commissions than did the nonequity agents. Additionally, only the equity agents handled "house" accounts, which (1) either originated in another agency that the company subsequently acquired or had originally been handled by an agent who had retired, and (2) did not fit the portfolio or expertise of any nonequity agent.¹³² The defining characteristics of the customers in these house accounts strongly suggest that they were less sensitive than other agency customers to the level of service they received from their agent. This hypothesis was corroborated by the fact that house accounts were three times more likely than other accounts to pay their premiums directly to their insurer,

129. See *supra* note 116 and accompanying text.

130. See Wilder, *supra* note 37, at 1-3.

131. *Id.* at 6.

132. *Id.*

rather than to pay them through the agency, indicating disengagement with their insurance agent.¹³³ The study concluded that contingent commissions significantly impacted the recommendations that the equity agents gave to their less responsive consumers, finding that “the prospect of contingency fees [led] equity agents to increase the frequency with which they placed house accounts with insurers offering contingency fees by more than 50%.”¹³⁴ In other words, equity insurance agents matched customers with different insurers depending on customers’ responsiveness, steering less responsive customers toward insurers that paid the agency contingent commissions.¹³⁵

A third reason that mere disclosure and competition cannot resolve many trilateral dilemmas is that the relationship between market intermediaries and their consumers often induces trust and reliance that limit consumers’ willingness to question the advice they receive.¹³⁶ Once again, this concern is significant in the insurance context. Many consumers develop longstanding relationships with their agents or brokers that induce both reliance and trust. Indeed, because of the complexities of insurance, an intermediary’s capacity to recommend proper coverage often critically depends on how well the intermediary understands the consumer’s particular needs.¹³⁷ Independent intermediaries therefore devote significant resources toward cultivating the trust of their clients.¹³⁸ As a result, intermediaries can sugarcoat potential deficiencies in insurance coverage and may avoid having to deal with them at all.¹³⁹ Of the few customers who inquire about their prospective insurer’s financial strength, virtually none will challenge their intermediary’s cursory response that the insurer’s capacity to pay claims is good—regardless of whether they are aware of their intermediary’s potential conflict of interest. This is particularly true given that insurance consumers must make numerous decisions about unfamiliar matters, such as deductibles, policy limits, endorsement options, and alterative insurers.¹⁴⁰

133. *See id.* at 7-8.

134. *Id.* at 10.

135. Although equity agents also tended to place a disproportionate percentage of their customers with insurers that maintained contingent commission contracts with the agency, the statistical significance of this effect was unclear. *Id.*

136. This concern, for instance, was one of the motivating factors that led Congress to regulate the trilateral dilemma involving real estate intermediaries in RESPA. *See Jackson & Burlingame, supra* note 25.

137. *See infra* notes 183-184 and accompanying text.

138. *See supra* notes 46-51 and accompanying text.

139. *See Tom Baker, Constructing the Insurance Relationship: Sales Stories, Claims Stories, and Insurance Contract Damages*, 72 TEX. L. REV. 1395, 1400-16 (1994).

140. Although any meaningful disclosure of contingent commissions would presumably occur before discussion of these matters, few consumers will literally walk out the door of a broker or agent upon seeing a disclosure form. Rather, to the extent that a disclosure can influence a customer’s decision about purchasing insurance through an intermediary, it is likely to do so after the customer has gathered basic information about the insurance options available through the intermediary.

Beyond Disclosure: The Case for Banning Contingent Commissions

Competition, even when it is facilitated by a disclosure requirement, is thus ultimately limited in its capacity to protect insurance consumers from being steered to inappropriate or overly expensive insurance. Although contingent commissions represent only a small portion of intermediaries' total compensation,¹⁴¹ they can nonetheless tempt intermediaries to provide biased advice to their customers, as the Arizona insurance agency study demonstrates. Competition can do little to check this temptation, even when the potential risk of steering is disclosed, because consumers have a limited capacity to identify biased advice from their intermediary. And when consumers do spot trouble, intermediaries can quickly change their tune, thus retaining the business of sophisticated insureds while continuing to exploit less sophisticated customers. Often not even this adjustment will be necessary, as intermediaries will be able to overcome consumers' potential concerns by invoking the relationship of trust that exists between them.

B. *Evaluating the Adverse Selection Argument*

The primary defense of contingent commissions is that they may reduce adverse selection by inducing intermediaries to convey risk information to insurers that the latter would otherwise be unable to observe.¹⁴² Although theoretically sound, this argument suffers from three significant limitations when it comes to consumer insurance markets. First, while adverse selection is a commonly theorized problem, the available empirical evidence suggests that it is not nearly as significant a problem in consumer insurance markets as the wealth of theoretical literature suggests. As a recent article by Professor Peter Siegelman explores in detail,¹⁴³ empirical studies on long-term care insurance,¹⁴⁴ term life insurance,¹⁴⁵ health insurance,¹⁴⁶ and automobile insurance¹⁴⁷ have each found no evidence of adverse selection.¹⁴⁸ There are

141. See *supra* note 57 and accompanying text.

142. See *supra* Section I.B.

143. See generally Siegelman, *supra* note 28.

144. Amy Finkelstein & Kathleen McGarry, *Multiple Dimensions of Private Information: Evidence from the Long-Term Care Insurance Market*, 96 AM. ECON. REV. 938 (2006).

145. Mark V. Pauly et al., *Price Elasticity of Demand for Term Life Insurance and Adverse Selection* (Nat'l Bureau of Econ. Research, Working Paper No. 9925, 2003), available at <http://papers.nber.org/papers/w9925.pdf>.

146. James H. Cardon & Igal Hendel, *Asymmetric Information in Health Insurance: Evidence from the National Medical Expenditure Survey*, 32 RAND J. ECON. 408 (2001).

147. Georges Dionne et al., *Testing for Evidence of Adverse Selection in the Automobile Insurance Market: A Comment*, 109 J. POL. ECON. 444 (2001).

148. The most compelling exception to this trend is a famous article documenting adverse selection in a health insurance market in which employee-insureds were able to select between alternative health insurance plans offered as a part of a subsidized group plan through an employer. David M. Cutler & Richard J. Zeckhauser, *Adverse Selection in Health Insurance*, 1 FRONTIERS HEALTH POL'Y RES. 1 (1998). But the risks of adverse selection in such "cafeteria style" insurance arrangements are unique. Thomas Buchmueller & John DiNardo, *Did Community Rating Induce an Adverse Selection Death Spiral? Evidence from New York, Pennsylvania, and Connecticut*, 92 AM. ECON. REV. 280, 280-81

various potential explanations for the persistent finding that most consumer insurance markets do not suffer from adverse selection. In many cases, such as life insurance, it is likely that insureds often do not have better information about their risk levels than insurers once contractually required disclosures by insureds are taken into account.¹⁴⁹ In other instances, such as automobile insurance, insureds may not be able to translate whatever informational advantage they have into better predictions about their risk levels than those generated by expert insurers.¹⁵⁰ This is particularly likely given emerging evidence that lay people are more subject to systematic cognitive biases than experts.¹⁵¹ A third explanation for the lack of adverse selection in most consumer insurance markets is that low-risk individuals also tend to be relatively risk-averse. This risk aversion leads low-risk individuals to value insurance more highly than high-risk individuals, which tends to counteract adverse selection.¹⁵² Whatever the explanation, though, the evidence seems clear that adverse selection is not nearly as significant a threat as defenders of contingent commissions suppose.

The second significant limitation to the adverse selection defense of contingent commissions is that it is unclear whether intermediaries who receive such commissions actually transmit better information to insurers than do other intermediaries. There is significant sociological evidence that captive agents, who work solely for a given insurer and thus do not receive contingent commissions, play a nontrivial underwriting role for those insurers.¹⁵³ However, the evidence that independent intermediaries also play this underwriting role when they receive contingent commissions is much less clear. The likelihood of such complicated information transfer may be low when it comes to the relatively run-of-the-mill underwriting decisions that underlie most consumer insurance markets.¹⁵⁴ Such underwriting requires simple and easily administrable algorithms,¹⁵⁵ which may well preclude

(2002). Moreover, because group insurance arrangements do not involve independent intermediaries at the level of consumers, contingent commission payments do not, of course, have any prospect of improving adverse selection in these markets.

149. Siegelman, *supra* note 28, at 1241.

150. *See id.* at 1242-47.

151. Nonexperts tend to place too much importance on salient events, to be overly optimistic in assessing their own risk levels, to treat low risk-levels as equivalent to zero, and to over-estimate the likelihood of affect-laden risks, such as that associated with nuclear waste. *See* Cass Sunstein, *The Laws of Fear*, 115 HARV. L. REV. 1119 (2002) (reviewing PAUL SLOVIC, *THE PERCEPTION OF RISK* (2000)). Experts, by contrast, can avoid such errors by focusing on a defined set of limited variables and by employing methods that tend to debias their results. *See id.* at 1144-56.

152. *See* Siegelman, *supra* note 28, at 1264-74.

153. *See supra* note 64.

154. *See supra* notes 171-173 and accompanying text.

155. *See* KENNETH ABRAHAM, *DISTRIBUTING RISK* 78 (1986) (“[A]n efficient classification system does not strive to make its premiums equal expected cost beyond the point where that goal is worth achieving.”).

Beyond Disclosure: The Case for Banning Contingent Commissions

incorporating nuanced information conveyed by intermediaries.¹⁵⁶

Finally, to the extent that insurers do indeed use information from intermediaries to make underwriting decisions, contingent commissions may not improve either the quality or the reliability of this information. An insurance company can generally monitor the type of insureds an intermediary refers to it merely by evaluating the loss ratio—a measure of insurer profitability¹⁵⁷—of the intermediary’s book of business. Insurers maintain loss ratios for each of their independent intermediaries in the ordinary course of business, and they can easily use those measures to determine whether an intermediary’s book of business generates more or less profit than the initial underwriting criteria suggested it would. Insurers can adjust their willingness to take insureds from particular intermediaries accordingly. Consequently, although contingent commissions may indirectly create an incentive for intermediaries to keep their loss ratios at a profitable level,¹⁵⁸ they are hardly necessary for this purpose given the more direct methods available to insurers—dropping, or threatening to drop, intermediaries with a consistently poor loss ratio. And unlike the defense of contingent commissions, this monitoring is not simply theoretical.¹⁵⁹ As one independent insurance agency operator explained, his agency “*guarded* [its] loss ratio very, very carefully because having a good loss ratio over the book of business is what enables you to negotiate favorable rates and go to new insurers and sign them up if you need them.”¹⁶⁰ Most insurers, the agent explained, carefully consider the loss ratios in agencies’ books of business when deciding whether to do business with them.¹⁶¹

C. *Toward Banning Contingent Commissions*

In consumer insurance markets, contingent commissions pose serious conflicts of interest for ostensibly independent intermediaries. They encourage these intermediaries to pursue higher commission rates rather than guide their

156. For this reason, “[i]ndividual companies are increasingly less likely to undertake their own home inspections or direct field investigation of an applicant. Instead, moral risk assessment is centralizing into data systems operated by information service companies that supply the insurance industry.” See ERICKSON ET AL., *supra* note 64, at 241.

157. See *supra* note 37.

158. See *supra* notes 68-71 and accompanying text.

159. Indeed, one underwriting manager for a multinational insurer has stated that “we have expressed to [brokers] time and again and built up over time what our likes and dislikes are from an underwriting standpoint.” If the broker can demonstrate that its “risks fit,” then the insurer “pursue[s] it aggressively . . . [.] cut[ting] incentive commission deals” for the broker. ERICKSON ET AL., *supra* note 64, at 227. Similarly, another underwriter explained that costs are often too high for insurers to investigate each potential insured on their own, and they often, therefore, need to “take a little bit on trust . . . dealing with the right broker.” *Id.* at 241.

160. *Id.* at 247.

161. *Id.* at 248.

customers to the insurance that best meets their needs or is most affordable. This conflict of interest is particularly problematic in insurance markets because consumers have an inherently limited understanding of their insurance needs and options. They often cannot, therefore, make optimal purchasing decisions without unbiased advice. Indeed, independent intermediaries justify their very existence on this basis, promising to look after their customers' needs. Given this fundamental information asymmetry, as well as the dynamics of the intermediary-customer relationship, disclosing contingent commissions to consumers will ultimately do little to mitigate the conflicts of interest that contingent commissions generate. Indeed, lawmakers have consistently recognized as much in the case of structurally similar trilateral dilemmas that plague other industries.¹⁶²

Meanwhile, the core defense of contingent commissions—that they reduce adverse selection by aligning the incentives of intermediaries and insurers—is largely illusory and is based on suspect assumptions when it comes to consumer insurance markets. Adverse selection is simply not a significant problem in consumer insurance markets. Even if it were shown to be a genuine problem, the claim that contingent commissions can actually alleviate the problem is both empirically unsupported and logically questionable given the limited underwriting criteria that most insurers use to price relatively small risks.

Simple cost-benefit analysis therefore strongly suggests that contingent commissions ultimately harm insurance consumers. Thus, there is a strong case to be made that state legislatures and regulators should go beyond a simple disclosure requirement to address these harms in consumer insurance markets. One intriguing, but ultimately limited, option is to replace contingent commissions with “supplemental compensation schemes” that base intermediaries' commission levels on the volume and profitability of business

162. As is described above, lawmakers have often gone beyond mere disclosure when addressing trilateral dilemmas in other contexts. *See supra* notes 102-114 and accompanying text. Admittedly, though, regulators have not always adopted an outright ban of side payments when faced with a trilateral dilemma. Section 8 of RESPA, for instance, only prohibits some types of side payments to real estate intermediaries. *See supra* notes 108-109 and accompanying text. Similarly, section 28(e) allows brokerage firms to make side payments to investment managers in the form of investment-related services. *See supra* note 111 and accompanying text. In part, these and other moderated responses to trilateral dilemmas may simply have been the result of political compromise. But they also may represent wise policy decisions that are driven by factors not present in the contingent commission context. For instance, some commentators have argued that allowing brokerage firms to pay investment managers for investment-related services helps to improve securities markets by encouraging managers to engage in efficient investment research. *See, e.g.,* Johnsen, *supra* note 111, at 91-104. Without these payments, the argument goes, investment managers would engage in too little investment research for a variety of reasons, including the fact that such research redounds to the benefit of all the investment manager's clients, creating a collective good problem. *See id.* at 95. Regardless of the persuasiveness of this argument, there is obviously little that is analogous in the case of contingent commissions. Ultimately, then, although most trilateral dilemmas tend to be resistant to disclosure-based solutions, the policy prescription in such cases depends on the specific features of the trilateral dilemma at issue.

Beyond Disclosure: The Case for Banning Contingent Commissions

that they have steered toward an insurer *in the past*. Although some have suggested that these newly emerging arrangements avoid the conflicts of interest associated with contingent commissions because they are “retrospective rather than prospective,”¹⁶³ this argument is clearly unpersuasive: the retrospectiveness of supplemental compensation arrangements merely shifts forward the potential pay-off of steering customers to suboptimal insurance. Instead of receiving increased contingent commissions in the current pay period for directing customers to a particular insurer, independent intermediaries will receive increased supplemental compensation in the next pay period for doing so.¹⁶⁴ Nonetheless, supplemental compensation arrangements, because they are set prior to the sale of an insurance product, may be easier to disclose to consumers than contingent commissions: unlike contingent commissions, the size of which are unknowable at the time of insurance sales, supplemental compensation arrangements provide intermediaries with a specific and quantifiable bonus for each additional consumer they direct to the insurer. They may be preferable to contingent commissions for that reason.

However, even retrospective supplemental compensation schemes coupled with a strong disclosure requirement are unlikely to limit meaningfully the risks of inefficient steering. Such schemes do not address the underlying reasons why disclosure has a limited capacity to protect consumers from trilateral dilemmas.¹⁶⁵ Regardless of how well consumers understand that their intermediaries face a potential conflict of interest, consumers are still subject to inefficient steering because they do not understand *the quality and price tradeoffs of their insurance options*. As such, most consumers will be unable to police the quality of the advice they receive from their intermediaries or to calculate the expected costs of this limitation. Those that are sophisticated enough to understand these issues will be unable to protect the interests of

163. Roberts, *supra* note 16 (quoting Robert A. Rusbuldt, CEO of Independent Insurance Agents and Brokers of America).

164. Moreover, supplemental compensation schemes would still produce variability in the commission rates that insurers pay because they link compensation to an individual producer’s “performance” for different insurers. So long as insurers vary their commission rates by the volume of business directed to it, differentials in commission rates will persist, and so too will the risk of steering. By contrast, when an insurer offers the same commissions to all intermediaries, regardless of the volume or profitability of their business, then competition should ensure that commission levels across insurers remain even for a given type of coverage. Indeed, anecdotal evidence suggests that premium commissions tend to be quite flat for a given type of insurance, though rigorous evidence on this point is not currently available. See *Money on the Table*, *supra* note 46 (quoting a number of experts who implicitly agree that premium commissions for a given line of insurance are currently quite flat due to competition, but who disagree about whether this result would obtain were contingent commissions banned). Although the size of premium commissions does tend to vary significantly across lines of insurance, see Cummins & Doherty, *supra* note 2, at 377-78 tbl.5 (collecting data from the National Association of Insurance Commissioners), that is hardly surprising, as the role of the intermediary depends significantly on the line of insurance in question.

165. Indeed, it is for precisely these reasons that the discussion in Section II.A explicitly assumed that consumers would read and understand a contingent commission disclosure. See *supra* note 119 and accompanying text.

other, less sophisticated, consumers, given the capacity of intermediaries to distinguish between sophisticated and unsophisticated insureds. And while an effective disclosure might constrain the relationship of trust that often exists between intermediaries and their clients, when it fails to do so consumers will still be subject to inefficient steering.

Given the inability of creative regulatory restrictions to limit the potential for inefficient steering, it is likely that the optimal regulatory approach is to ban contingent commissions in consumer insurance markets.¹⁶⁶ Indeed, the cost-benefit analysis described above suggests that the only clear drawback to this proposal is the administrative cost of implementing the rule. Although the mere act of banning contingent commissions would not be costly, policing that ban might be. In particular, insurers could replicate contingent commission arrangements with retrospective supplemental compensation schemes. This commission arrangement would be significantly more difficult for regulators to police because it could be accomplished without the insurers ever publicly revealing their rationale for offering variable commission rates. Ultimately, though, a sufficiently clear statute or regulation, combined with a large penalty, should dissuade insurers from adopting this type of approach.¹⁶⁷ In the final calculus, then, the benefits of banning contingent commissions and similar insurer-financed bonuses for independent intermediaries—at least in consumer insurance markets—are likely to significantly exceed the costs.

III. CONTINGENT COMMISSIONS IN INSURANCE MARKETS WITH SOPHISTICATED INSURED

Although disclosure may be an insufficient response to contingent commissions in consumer insurance markets, it is less clear whether something more than disclosure is needed in commercial insurance markets. This Part explores the likely impact of disclosure in these markets, where insurance purchasers are far more sophisticated than conventional consumers. It concludes that there may be sound reasons to consider banning contingent commissions in the commercial context, but that more empirical research is needed before a regulatory approach can be confidently embraced.

Section A begins the argument by showing that even if disclosure

166. The inability of supplemental compensation schemes to limit the pernicious effects of contingent commissions should be contrasted with the soft-money compromise that allows brokerage firms to pay investment managers for investment-related services. This partial ban, at least in theory, helps to improve securities markets by encouraging investment managers to engage in efficient investment research. See Johnsen, *supra* note 111, at 91-104; see also *supra* note 162. As is described above, however, no one has yet proposed a similar partial ban of contingent commissions that could effectively limit the risk of inefficient steering created by contingent commissions.

167. Another potential strategy for implementing a ban would be to require insurers to pay a single, market-determined, premium commission rate for a given policy type to all independent intermediaries with whom they do business.

Beyond Disclosure: The Case for Banning Contingent Commissions

meaningfully informs insurance purchasers about the costs of contingent commissions, market outcomes cannot necessarily be trusted to reflect actual purchaser preferences. The intuition behind this outcome is that insurance purchasers who understand the risks of contingent commissions may rationally choose a market outcome that is inconsistent with their true policy preferences in order to signal to insurers that they are low-risk. Section B extends this argument to show that the contingent commissions may have a limited capacity to mitigate information asymmetries in sophisticated insurance markets, in which adverse selection may indeed be a significant problem. This Section presents an alternative model of contingent commissions in which insurance purchasers are allowed to mimic low-risk insureds when dealing with insurance intermediaries. Under the model, the benefits of contingent commissions in reducing adverse selection are smaller than earlier models suggest, and they are potentially swamped by new, additional costs resulting from high-risk insureds engaging in socially inefficient strategic behavior.

A. *Steering and Signaling in Sophisticated Insurance Markets*

Even in insurance markets where insurance purchasers tend to be relatively sophisticated and informed, disclosure of contingent commissions may fail to cause purchasers to guard sufficiently against the risk of steering. The basic reason is that informed and rational insureds who choose to purchase insurance through non-contingent-commission intermediaries may be penalized for this decision, as their choice signals to insurers that they are likely to be higher-risk than their directly observable characteristics suggest.¹⁶⁸ As defenders of contingent commissions have argued,¹⁶⁹ intermediaries who accept contingent commissions may have an incentive to help insurers correctly set prices by providing them with additional information about insureds' risk levels. But while accurate pricing improves the efficiency of the insurance market as a whole, it harms insurance purchasers who are higher-risk than insurers would otherwise believe. As such, these insureds will seek to hide their high-risk status. One way of doing so is by purchasing insurance through intermediaries that do not use contingent commissions—and that therefore do not communicate additional information to insurers.¹⁷⁰ Given this general

168. The assumption here that purchasers are “informed and rational” means that they understand that contingent commissions may induce intermediaries to reveal information to insurers. Arguably, there is an intermediate stage in which purchasers would not understand this implication of contingent commissions but would be able to determine that the costs of steering are relatively small and make their decisions accordingly. However, even this intermediate stage of information would still make it difficult to rely on insurance purchasers' market decisions as expressing their actual preferences. Without understanding that contingent commissions improve communication between insurers and intermediaries, purchasers would have little sense of the potential *benefits* of contingent commissions.

169. See *supra* Section I.B.

170. A second way for insurance purchasers who pose a higher risk than their clearly observable

preference, insurers will tend to charge those who purchase insurance through such intermediaries a higher premium than their observable characteristics would suggest should be charged.

This result is counterintuitive because, in most markets, effective disclosure to sophisticated purchasers—whether the underlying risk concerns the potential safety failure of a product or the prospect that a service will be ineffective—is enough to ensure that regulatory oversight is not needed.¹⁷¹ Insurance markets are different, however, because the purchaser's attributes largely define the product: the meaning of a promise to pay for specified losses is dependent on the promisee's likelihood of suffering those losses. For this reason, even if insurance purchasers are fully informed about contingent commissions (through disclosure or otherwise), the market cannot be relied upon to produce the socially optimal result.

A key assumption in this argument—as in the adverse selection defense of contingent commissions more generally¹⁷²—is that insurance intermediaries can and do communicate insureds' risk levels to insurers when they receive contingent commissions, but not otherwise. Although this claim is questionable when it comes to consumer insurance markets,¹⁷³ it is more plausible in commercial insurance markets where the insureds are sophisticated actors. Underwriting decisions in such insurance markets tend to be less mechanistic than in consumer markets because they involve more risk, higher premiums, and more heterogeneous and sophisticated insureds.¹⁷⁴ This lack of mechanistic underwriting, combined with lower numbers of insureds per intermediary, may also limit the capacity of insurers to make inferences about intermediaries based on their loss ratios.¹⁷⁵ In any event, the assumption that contingent commissions improve insurer information plays just as vital a role here as it does in the claim that contingent commissions can reduce adverse selection.¹⁷⁶ To the extent that the assumption is incorrect, both arguments fail. Yet if

characteristics might suggest to hide this fact is to mimic low-risk insureds when dealing with their insurance agents or brokers. See *infra* Section III.B. The result described in this Section, whereby insureds who purchase insurance through non-contingent-commission intermediaries pay a market penalty, still applies under that alternative assumption.

171. If consumers understand the risks associated with a product or service and nonetheless choose to purchase it, then they generally must value the product or service more than the combined cost of the purchase price and expected risk. See, e.g., STEVEN SHAVELL, *ECONOMIC ANALYSIS OF ACCIDENT LAW* 51-53 (1987) (arguing that products liability law is not needed if consumers are informed about product risks); Alan Schwartz & Louis L. Wilde, *Imperfect Information in Markets for Contract Terms: The Examples of Warranties and Security Interests*, 69 VA. L. REV. 1387 (1983) (arguing that courts should not intervene in standard form contracts if a sufficiently large minority of consumers are informed about these contracts).

172. See *supra* Section I.B.

173. See *supra* notes 153-160 and accompanying text.

174. Cf. *supra* text accompanying notes 154-156.

175. Cf. *supra* text accompanying notes 157-161.

176. See *supra* Section I.B.

Beyond Disclosure: The Case for Banning Contingent Commissions

contingent commissions do not reduce adverse selection, then the case for banning them is clear: there is little reason to permit a costly practice that has no benefits.

To develop this signaling argument more thoroughly, consider a group of insureds that, from an insurer's perspective, all have the same observable risk level, designated R_O . Suppose that each insured also has one of two actual risk levels, which are equally distributed among the population.¹⁷⁷ Some insureds, designated R_L , have a lower actual risk level than their observable characteristics suggest, and some insureds, R_H , are actually more risky than their observable characteristics indicate. For ease of discussion, let us define the difference between insureds' actual risk levels and their observed risk levels as their "residual risk." Thus, R_L types have a negative residual risk and R_H types have a positive residual risk. Both R_L and R_H know their own risk types. Suppose further that insureds can purchase insurance through two types of intermediaries,¹⁷⁸ both of which are capable of identifying whether an insurance purchaser is R_H or R_L . Some intermediaries, designated CC , receive contingent commissions and, for this reason, will credibly communicate an insured's residual risk level to insurers. Doing so allows the intermediary to maximize contingent commissions because it facilitates the appropriate pricing of policies. Other intermediaries, designated NCC , do not receive contingent commissions and therefore do not contribute to the insurer's underwriting efforts by transmitting information about potential insureds' residual risk.¹⁷⁹

This simple model helps illustrate why, if informed purchasers are allowed to choose between CC and NCC intermediaries, insurers will rationally conclude that those who choose an NCC intermediary are more likely that not to have a high residual risk (R_H). R_H types will generally have an incentive to hide their residual risk from insurers so that they are not charged an actuarially fair premium, which is higher than they would otherwise pay. They cannot do so by going to a CC intermediary, as the CC intermediary conveys this risk

177. This is not nearly as strong an assumption as it first appears to be. Insurers are generally thought to be quite good at calculating roughly accurate averages of risk levels across defined populations. If the average risk level is close to the median, then there will be an approximately equal number of people who are more risky than average compared to those who are less risky than average.

178. This assumption can be challenged because purchasers can bypass intermediaries altogether and obtain insurance through an insurer that offers direct underwriting, either over the phone or the Internet. ABRAHAM, *supra* note 13, at 56. But this option, of course, requires completely foregoing the advice of an intermediary who can identify insurance needs and suggest various options. Moreover, it is only available from a limited number of insurers for a limited number of insurance lines, such as automobile and homeowners insurance. *See id.*

179. Another objection is that under some reform proposals, insureds can force an agency not to accept contingent commissions. In this scenario, a single agent or broker would accept contingent commissions or not depending on each client's stated preference. *See Fitzpatrick, supra* note 3, at 3069-70. This scenario, however, would not impact the analysis because insurers, of course, would know in any given case whether an agent or broker was accepting contingent commissions. They could therefore simply treat the intermediary differently based on whether or not it was accepting contingent commissions for a given transaction.

information to the insurer.¹⁸⁰ By going to the *NCC* intermediary, by contrast, R_H types can attempt to hide their true status. Contrast this with R_L types, who actually gain from credibly communicating their true risk status to insurers, and who therefore have an affirmative incentive to purchase insurance from a *CC* intermediary. Countervailing factors, most notably the risk of receiving biased advice due to steering, may, of course, counteract this effect. But so long as any countervailing forces equally affect both R_H and R_L types, insurers can rationally infer that a higher percentage of R_H than R_L types will prefer *NCC* intermediaries, and a higher percentage of R_L than R_H types will prefer *CC* intermediaries.¹⁸¹ Insurers, knowing this, will set their premiums accordingly and will impose a surcharge for the purchase of insurance through an *NCC* intermediary.

This “surcharge” decreases the extent to which informed purchaser choices reveal actual purchaser preferences. The reason is that the cost insurers add to purchases through an *NCC* intermediary is not tethered to the actual costs associated with that choice. To see why this matters, consider a low residual risk insured (R_L) who, if given the choice between banning contingent commissions and mandating them, would prefer a ban because of the possible costs associated with steering. Even though this insured is R_L , he or she would prefer to be lumped together with all insureds and treated as having an average risk level (R_O) rather than pay the costs associated with steering and be recognized as a genuine R_L . But when given the choice between these two options in the marketplace, this insured might nonetheless choose to purchase insurance from a *CC* intermediary because purchasing from an *NCC* intermediary would falsely signal that the insured had a high residual risk level, resulting in the insured paying a higher premium than would be paid by an average risk (R_O).

B. *Mimicry: A New Adverse Selection Model for Sophisticated Insurance Markets*

As noted above, the core defense of contingent commissions—that they reduce adverse selection—has some traction in sophisticated insurance markets, where there are fewer reasons to be skeptical about the importance of adverse selection.¹⁸² But even in these sophisticated insurance markets, contingent commissions may fail to improve adverse selection problems

180. This is not true, of course, if R_H types can hide their risk status when dealing with *CC* intermediaries. The implications of such deception are considered *infra* Section III.B.

181. In theory, if the costs of steering were sufficiently high for all insureds, then both R_H and R_L types would go to *NCC* intermediaries. This result is unlikely given that different insureds are likely to weigh the costs of steering differently. In any event, this result would be the functional equivalent of banning contingent commissions, and so does not undermine the analysis.

182. See *supra* notes 143-152 and accompanying text.

Beyond Disclosure: The Case for Banning Contingent Commissions

significantly. To make this argument, this Section expands on the model of contingent commission payments presented above, but adds an additional assumption—that R_H insureds can, for a cost, mimic R_L insureds when interacting with intermediaries. This assumption is first explained and justified in Subsection 1. The model, which is developed in Subsection 2, suggests that R_H insureds may choose to purchase insurance through CC intermediaries and to mimic R_L insureds strategically in a way that both creates new social costs and mitigates the capacity of contingent commissions to reduce adverse selection. Under the model, contingent commissions shift, rather than eliminate, the informational barrier that causes adverse selection. In other words, while contingent commissions encourage the free flow of information between insurers and intermediaries, they simultaneously discourage the free flow of information between insurance purchasers and intermediaries.

1. *The Mimicry Assumption*

Sophisticated insurance purchasers ordinarily have a clear incentive to be honest with their insurance brokers or independent agents. Independent insurance intermediaries help “buyers to identify their coverage and risk management needs and [match] buyers with appropriate insurers.”¹⁸³ These risk-management services are particularly important for sophisticated insurance purchasers, who tend to have complicated risk-management needs. If these potential insureds are not truthful and forthright about their potential liabilities, then their intermediaries will not be capable of providing these services. An intermediary’s capacity to identify an insured’s vulnerabilities and to offer concrete methods for reducing risk is obviously compromised when insureds withhold complete and accurate risk information. Similarly, if insureds are not completely forthright about their risk profiles, their agents or brokers may be unable to recommend or investigate truly comprehensive insurance options. In fact, many endorsement options are specifically designed so that they will be particularly appealing to relatively high-risk insureds.¹⁸⁴ Intermediaries that do not know that an insured is relatively high-risk will have little reason to investigate or recommend these endorsements.

When intermediaries receive contingent commissions, however, insurance purchasers’ ordinary incentive to be truthful with their brokers or agents may be overwhelmed by a contrary incentive. In the model outlined above, informed and sophisticated insurance purchasers will understand that an agent or broker who receives contingent commissions may disclose otherwise unobservable

183. Cummins & Doherty, *supra* note 2, at 360.

184. This is the classical solution to the adverse selection problem pioneered by Rothschild and Stiglitz. See Rothschild & Stiglitz, *supra* note 72, at 643.

risk information to insurers.¹⁸⁵ For this reason, informed R_H insureds may have an incentive to hide their residual risk level when they purchase insurance through a CC intermediary. By mimicking an R_L insured and purchasing insurance through a CC intermediary, the high residual risk insured may be able to receive a discounted premium from the insurer: the CC intermediary, believing the customer to be R_L , will signal this misunderstanding to the insurer, who will charge the incognito R_H customer a low premium.

In many cases, determined R_H insurance purchasers will be able to successfully mimic R_L insureds when interacting with intermediaries. To do so, the insured need not alter verifiable or easily observable facts—any such information is presumably observable by insurers directly and is thus not part of the insured's residual risk.¹⁸⁶ Rather, the insured would merely need to alter evidence relevant to residual risk, meaning evidence that the insurer would not observe without the input of the intermediary. For instance, the insurance purchaser might fail to mention idiosyncratic risks about which the producer did not directly inquire. The normally nonchalant businessman might well put on a front of hyper-vigilance for his visits to his broker. If this proved impossible, he might send a particularly safety-minded subordinate to meet with the broker or agent. In cases in which a broker or agent had historical knowledge of the company's poor safety habits, the insured could switch intermediaries and wipe the slate clean. Such deception is likely to be successful because an intermediary's comparatively superior knowledge about an insured depends almost exclusively on his or her direct interactions with the insured.¹⁸⁷

Of course, these efforts at mimicry would be costly for insurance purchasers. First, and most obviously, mimicry would require the additional effort of strategizing about how to mimic an R_L insured. Second, insureds who did mimic would forego at least some of the risk mitigation advice that independent intermediaries provide to their clients. Third, as noted above, insureds who camouflage their risk levels potentially forego receiving advice about desirable endorsements or alternative policy options.

2. *The Consequences of Mimicry*

To understand the circumstances under which R_H insureds will rationally choose to mimic R_L insureds and the consequences of such mimicry, consider

185. To the extent that even sophisticated insurance purchasers do not understand this, there is even more reason to believe that steering is a significant problem and that contingent commissions genuinely decrease the quality of insurance that these purchasers ultimately receive. See *supra* Section II.B (discussing how, even assuming that insureds understand the implications of an intermediary's contingent commission disclosure, they still may be steered to inefficient coverage).

186. See *supra* note 177 and accompanying text.

187. See *supra* notes 64-67 and accompanying text.

Beyond Disclosure: The Case for Banning Contingent Commissions

the following extensions of the model introduced in Section A. Suppose that the actuarially fair premium for all insureds, when insurers cannot distinguish between residual risk types, is P . For R_L insureds, the actuarially fair premium is reduced by an amount D (for a total of $P-D$) and for R_H insureds, it is increased by D (for a total of $P+D$).¹⁸⁸ Insureds that purchase through a *CC* intermediary must pay a cost, S , in addition to their premium, which represents the potential that they will be “steered” to a low-quality insurance arrangement for a given price. This might include relatively less comprehensive coverage, a less financially secure insurer, or an insurer with a bad reputation for claims payment.¹⁸⁹ Because insureds have variable insurance needs and information about those needs, the size of S is assumed to vary among both R_H and R_L insureds.

Only *CC* intermediaries can credibly communicate information to insurers about an insured’s residual risk. R_H insureds, however, may not be forthright about their residual risk level, and they are capable of either revealing their true residual risk level to brokers or mimicking the risk level of an R_L insured. Intermediaries cannot distinguish between actual R_L insureds and R_H insureds who mimic low-risk status. To mimic an R_L insured, high-risk insureds must pay a cost X that reflects the costs of changed behavior, decreased risk mitigation advice, and the possibility of foregoing desirable endorsement or alternative policy options.¹⁹⁰ Again, because insureds will have different insurance needs and information about those needs, the cost of X will differ among R_H insureds. These basic parameters, including those introduced above,¹⁹¹ are summarized below:

$$R_L < R_O$$

$$R_H > R_O$$

$$\text{For } R_L, \text{ Premium} = P-D$$

$$\text{For } R_H, \text{ Premium} = P+D$$

$$\#R_L = \#R_H$$

CC: can communicate the insured’s residual risk level

NCC: cannot communicate the insured’s residual risk level

S : Increased steering cost an insured pays for a *CC* intermediary (varies among insureds)

X : Amount R_H must pay to mimic R_L (varies among R_H insureds)

Given these basic parameters, consider how insurers and purchasers will respond. First, note that *insurers will only set two premiums: one for insureds who purchase insurance through an *NCC* intermediary, and one for insureds*

188. As above, there are an equal number of R_H and R_L insureds.

189. Even if the insurance customer can prevent these outcomes through monitoring, such monitoring is presumably not free, and S can thus represent these costs. Additionally, S captures the theoretical possibility that producers will tend to be less aggressive in seeking payment for their clients’ claims when they receive contingent commission payments. *See supra* note 41.

190. *See supra* Subsection III.B.1.

191. *See supra* Section III.A.

who purchase insurance from a *CC* intermediary who reports that a purchaser appears to be R_L . Insurers will not set a third premium for insureds who purchase insurance from a *CC* intermediary reporting that a purchaser appears to be R_H . The reason is that R_H insureds will never reveal their true status to a *CC* intermediary. Doing so would lead the *CC* intermediary to tell the insurer about R_H 's high-risk status, causing the insured to pay the premium $P+D$. Instead, it will always be a better option (though not always the best), for the R_H insured to purchase through *NCC* intermediaries, who cannot communicate residual risk levels. Because insurers do not receive any credible information about the residual risk level from *NCC* intermediaries, they would not know that $P+D$ was the actuarially fair amount to charge. Even if most customers of *NCC* intermediaries were R_H —as Section III.A predicts—the insurer could not presume that they all were; at least some R_L purchasers would choose to purchase through an *NCC* broker if they viewed S to be large, and thus the actuarially fair premium for insurance purchased through an *NCC* broker would have to be less than $P+D$.

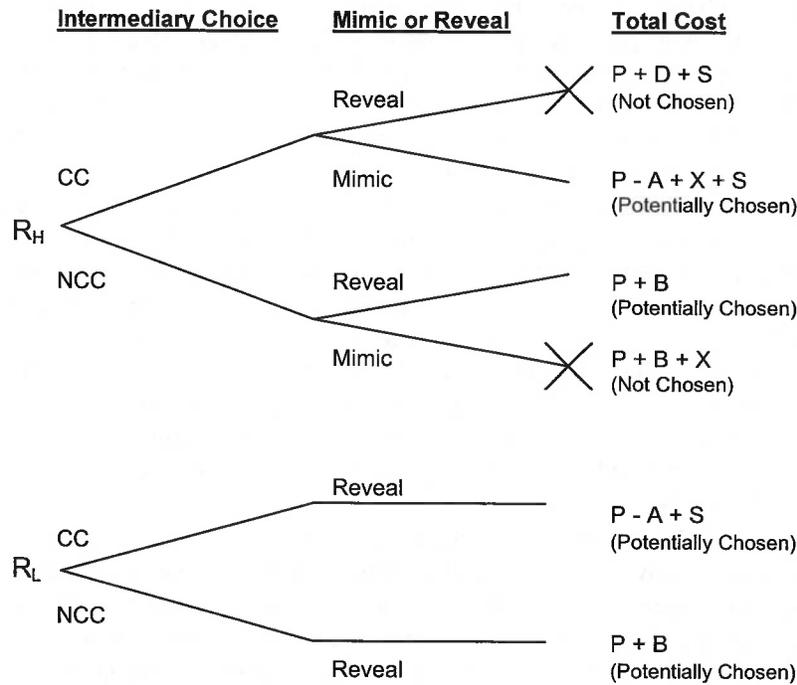
Second, note that *insurers, in setting premiums, can partially induce self-sorting by R_L and R_H types if they charge a lower premium to insureds who purchase through *CCs* than those who purchase through *NCCs**. Both R_L and R_H insureds will always prefer to purchase insurance from an *NCC* unless they can receive a lower premium through a *CC*: purchasing through a *CC* imposes cost S , and there is no countervailing benefit unless the price of insurance is lower.¹⁹² But insurers can induce some insureds to purchase insurance through a *CC* intermediary by charging a lower premium, which will offset the cost of S for some insureds. This group of insureds who purchase from a *CC* in response to a decrease in premiums will consist primarily of R_L rather than R_H types. The reason is straightforward: the total cost to R_L insureds of purchasing through a *CC* intermediary will be less than for R_H insureds because the R_H insureds must pay X , the cost of mimicking, if they purchase through a *CC*. (Recall from above that it will never be sensible for R_H to purchase from a *CC* and not to mimic.) By contrast, both R_L and R_H pay precisely the same cost when they purchase through an *NCC* intermediary, as that intermediary cannot credibly signal any residual risk information to the insurer.

Based on these conclusions, *the insurer's optimal pricing strategy will be to charge a premium $P-A$ (for some A such that $A < D$) for insureds who purchase through *CC* intermediaries, and to charge a premium $P+B$ (for some B such that $B < D$) for insureds who purchase through *NCC* intermediaries*. This strategy is superior to either charging the same price to all insureds, or charging a higher price to insureds who purchase through *CC* intermediaries, because

192. Notably, this is consistent with the possibility that contingent commissions reduce adverse selection. To the extent that a *CC* intermediary can reduce adverse selection, purchasers will experience this effect in a reduced premium.

Beyond Disclosure: The Case for Banning Contingent Commissions

neither of those strategies achieves any advantage in appropriately pricing risk. Moreover, the strategy is optimal when A and B are both less than D because the pool of insureds who purchase from either type of intermediary will almost certainly be heterogeneous. The only scenario in which the insurer could achieve a complete separating equilibrium¹⁹³ is if the premium discount for CC intermediaries over NCC intermediaries outweighed all R_L 's valuations of S , but did not outweigh any R_H 's valuations of $S+X$. Because different insureds value S and X differently under the model's assumptions, such a separating equilibrium does not occur. This analysis is presented visually, in the following figure:



- P = Premium for average risk insured
- D = Actuarially fair change in premium based on risk level
- X = Penalty for mimicry
- S = Steering cost
- B = Surcharge for purchasing from non-contingent commission broker
- A = Bonus for purchasing from contingent commission broker

193. In a separating equilibrium, all purchasers of different types choose different options, thus fully revealing their type indirectly. See DOUGLAS G. BAIRD ET AL., GAME THEORY AND THE LAW 140-47 (1994).

As this simple model clarifies, R_H insureds will tend to mimic R_L insureds if $P - A + X + S < P + B$. By simple arithmetic, this reduces to $X + S < B + A$. Thus, R_H insureds will mimic if the difference in premiums between CC and NCC intermediaries (i.e., $B + A$) is larger than the cost of steering (S) plus the cost of mimicry (X). When this is the case, R_H insureds will find it profitable to incur the costs of both S and X .

Moreover, *mimicry results in potentially substantial social costs and mitigates the capacity of contingent commissions to reduce adverse selection.* All of the costs of mimicry (X) and some of the costs of steering (S) are social as well as private costs: overall social welfare decreases when insureds undertake costly measures to change their behavior, forego valuable risk mitigation advice, chance purchasing insufficient insurance, or purchase insurance that is of poor quality. Simultaneously, the private benefit to R_H of mimicking R_L —the premium reduction of $B + A$ —is not a social benefit and actually limits the capacity of contingent commissions to mitigate adverse selection.¹⁹⁴ Indeed, under the model, R_H insureds never reveal their status to CC brokers or agents, meaning that these agents never have any information to reveal to insurers. Importantly, adverse selection does decrease in the model, but not for the reason highlighted by other writers or to the extent that these writers have suggested. Instead, adverse selection only decreases because insureds who purchase through a CC intermediary are more likely to be low-risk than insureds who purchase through an NCC intermediary, allowing insurers to address adverse selection by charging higher premiums to the latter group of insureds. But this method for reducing adverse selection is limited: it only allows insurers to improve their information imperfectly, by making potentially incorrect inferences about insureds' residual risk levels based on their purchasing decisions.

Ultimately, the model suggests that the costs of mimicry, *even ignoring the cost of steering*, may exceed the benefits of reduced adverse selection. Offering sophisticated purchasers a choice about how their intermediaries are compensated may induce high-risk insureds to take socially costly measures in order to transfer wealth from low-risk insureds to themselves. The opportunity to mimic a low-risk insured (and therefore to receive a lower premium) may be an enticing benefit for the high-risk insured, but it is actually a social cost.

194. Note that in the model, as observed *supra* Section III.A, insureds may well prefer a system in which contingent commissions are banned to one in which they have the choice of CC or NCC intermediaries. In a world without contingent commissions, low-risk insureds, like high-risk insureds, pay a premium P for insurance. Compare this with the equilibrium result in a world with disclosure: low-risk insureds have the choice of paying premium $P + B$ or $P - A + S$. If $S > A$, then low-risk insureds are unambiguously worse off in a world where contingent commissions are available as compared to a world in which such payments are banned. The same reasoning applies to high-risk insureds: following the same logic as above, they will be worse off if $X + S > A$. In other words, if low-risk insureds are worse off, then so too are high-risk insureds.

Beyond Disclosure: The Case for Banning Contingent Commissions

Indeed, it allows high-risk insureds partially to undo the benefits of reduced adverse selection that proponents of contingent commissions have identified. To accomplish this feat, high-risk insureds must incur a number of real costs associated with convincing their broker that they are low-risk insureds.

C. *A Possible Need for Intervention in Sophisticated Insurance Markets*

The preceding analysis suggests that, under a plausible set of assumptions, contingent commission payments may reduce the efficiency of commercial insurance markets, even though insurance purchasers in these markets are both informed and rational. First, it explains how even informed insurance purchasers may opt for the services of an intermediary that accepts contingent commissions in order to signal to insurers that they view themselves to be low-risk. The fact that these informed purchasers rely on contingent commission intermediaries does not, therefore, prove that this is the efficient market outcome. In fact, as Section B has explained, the costs of contingent commission payments in these markets are not limited to the prospect of steering. Sophisticated insurance purchasers, knowing that intermediaries who receive contingent commissions have an incentive to communicate risk information to insurers, may choose to purchase insurance through a broker or agent that accepts contingent commissions while masking information suggesting that they are relatively high-risk. High-risk insureds' "mimicry" of low-risk insureds not only undermines the capacity of contingent commissions to reduce adverse selection, but it creates additional social costs: it impedes the intermediary's ability to offer his or her clients risk-mitigation advice or to suggest appropriate coverage options.

The policy implications of these theoretical models, of course, are unclear and require empirical work testing the assumptions and predictions of the competing theoretical models. Two predictions, in particular, could be tested based on the theoretical model above. First, Section A has predicted that insureds who purchase through intermediaries that have disclaimed contingent commissions will pay premiums that are higher than what they otherwise would have paid. Second, Section B has predicted that high-risk insureds will be less forthcoming with contingent commission intermediaries than they are with intermediaries that disclaim contingent commissions. Because almost all intermediaries accepted contingent commission payments from some insurers until 2004, the possibility of testing these propositions has only recently emerged. Empirical work in this area is, therefore, both an exciting and promising frontier for research.

CONCLUSION

Attorney General (now Governor) Spitzer's opening salvoes against

contingent commissions held the promise of reshaping the relationship between insurers and independent intermediaries. In the last two years, however, insurers, brokers, independent agents, and academics have turned the tide, transforming accusations of fraud in core industry practices into a mere requirement that insurance consumers be shown one more standardized form. Unfortunately, such disclosure will do little to limit the risks of steering in consumer markets. Disclosure may alert consumers to the fact that their intermediaries have a potential conflict of interest, but it will not equip them to limit the consequences of this conflict of interest, in light of the inherent informational asymmetries existing between insurance consumers and intermediaries. Lawmakers must weigh these costs against the theoretical possibility that contingent commissions reduce adverse selection, a problem that empirical evidence suggests is insignificant in consumer insurance markets. The resulting cost-benefit analysis provides a strong case for going beyond disclosure in consumer insurance markets—and perhaps even banning contingent commission payments in all their manifestations.

By contrast, the optimal regulatory solution is unclear when it comes to sophisticated insurance markets, where insurance purchasers can meaningfully be informed about contingent commissions. There are, however, theoretical reasons to believe that even in these insurance markets, disclosure is too limited a response. Insurance purchasers might well make market choices that do not reflect their actual preferences regarding the desirability of contingent commissions. This result may exacerbate adverse selection and may create additional social costs. Although it is too early to embrace regulatory oversight of contingent commissions in sophisticated insurance markets, it is also premature to assume that market forces will produce efficient results.

90

**DIFFERENTIAL COMPENSATION AND THE
“RACE TO THE BOTTOM” IN
CONSUMER INSURANCE MARKETS**

*Daniel Schwarcz**

This contribution to a symposium on insurance intermediaries analyzes insurers' compensation of independent agents and brokers in consumer markets. It focuses on various forms of "differential compensation," whereby an intermediary's compensation differs depending on the insurer with which the consumer ultimately purchases coverage. Such differential compensation, the article argues, undermines competition among consumer insurers with respect to non-price product attributes. This, in turn, increases the risk of a "race to the bottom" in consumer insurance markets, as insurers focus on selling the cheapest coverage possible that is consistent with legal restrictions. To address these problems, this article suggests that insurers who rely on independent agents to sell consumer lines of insurance should be prohibited from paying different rates of compensation to different agents for the sale of the same line of insurance.

In 2004, a series of lawsuits filed by the New York Attorney General challenged insurers' long-standing payments of year-end bonuses to insurance brokers. The lawsuits alleged that these payments, known as contingent commissions, created conflicts of interest that undermined

* Associate Professor, University of Minnesota Law School. For helpful comments, I thank Hazel Beh, Tom Cotter, Sean Fitzpatrick, Kristin Hickman, Claire Hill, Brett McDonnell, Francesco Parisi, Jeffrey Stempel, an anonymous referee, attendees of the Insurance Intermediaries panel at the 2008 Annual Association of Law Schools, and participants in a research seminar at the Department of Risk Management and Insurance at Georgia State University. This symposium piece builds off of my earlier article, Daniel Schwarcz, *Beyond Disclosure: The Case for Banning Contingent Commissions*, 25 YALE L. & POL'Y REV. 289 (2007).

brokers' professed loyalty to their clients.¹ "If the practices identified in our suit are as widespread as they appear to be," the Attorney General stated, "then the industry's fundamental business model needs major corrective action and reform."²

Within months of these allegations, the commercial insurance industry had indeed changed significantly. Each of the four largest insurance brokers pledged to end their practice of accepting contingent commission payments from insurers.³ Because of the concentration of the insurance brokerage industry – the three largest brokers, Marsh, Aon, and Willis enjoyed more than a 54% market share among the top 100 brokers in 2004⁴ – this shift dramatically impacted the entire market. Meanwhile, the prominence of these allegations led corporate risk managers and other sophisticated insurance purchasers to demand from their brokers previously-undisclosed details about contingent commission arrangements.⁵ Although many small brokers still accept contingent commissions, many other brokers (including the four largest) now publicly tout their refusal to accept such commissions in marketing themselves to their clients.⁶

¹ See generally Sean M. Fitzpatrick, *The Small Laws: Eliot Spitzer and the Way to Insurance Market Reform*, 74 *FORDHAM L. REV.* 3041, 3047 (2006) (reviewing the trajectory of Attorney General Spitzer's investigation of contingent commissions); Daniel Schwarcz, *Beyond Disclosure: The Case for Banning Contingent Commissions*, 25 *YALE L. & POL'Y REV.* 289 (2007).

² Press Release, Eliot Spitzer, New York Attorney Gen., Investigation Reveals Widespread Corruption in Insurance Industry (Oct. 13, 2004), available at http://www.oag.state.ny.us/press/2004/oct/oct14a_04.html.

³ See Schwarcz, *Beyond Disclosure*, *supra* note 1, at 291-92.

⁴ See J. David Cummins & Neil A. Doherty, *The Economics of Insurance Intermediaries*, 73 *J. RISK & INS.* 359, 364 (2006).

⁵ See David Dwanka, *Mid-Level Insurance Brokers Defend Contingent Commissions Amid Growing Criticism*, *BESTWIRE*, May 8, 2006.

⁶ *Broker Compensation: Hearing Before the New York Ins. Comm'r* (2008) (testimony of Don Bailey, CEO of Willis) (testifying that Willis does not accept contingent commissions because they pose a "clear and obvious conflict of interest") available at http://www.ins.state.ny.us/agbrok/br_cmp_idx.htm.

For all of this reform in *commercial* insurance markets, virtually nothing has changed about the way intermediaries in *consumer* insurance markets are compensated. In both property/casualty and life/health consumer insurance lines, most independent insurance agents continue to receive increased compensation from insurers to whom they steer a significant amount of business.⁷ And, unlike sophisticated insurance purchasers, most consumers continue to have no real understanding of these practices and the impact they may have on the advice that insurance agents offer.

From a doctrinal perspective, this divergence in consumer and commercial insurance markets may appear to be perfectly reasonable. The insurance brokers that service commercial insurance markets are generally considered to be legal agents of policyholders.⁸ By contrast, the independent insurance agents that populate consumer insurance markets are usually described primarily as legal agents of insurers, rather than consumers, and therefore have more limited (if any) fiduciary obligations to policyholders.⁹ Consequently, compensation structures that create conflicts of interest appear to be more troubling doctrinally in commercial markets than in consumer markets.

But from an economic perspective, the differential reform in commercial and consumer insurance markets is bizarre. Unlike sophisticated commercial entities, ordinary consumers generally have limited information about the relative quality of different carriers and a bounded ability to translate the information they do have into effective decision-making.¹⁰ As Cass Sunstein and Richard Thaler recently observed, “the benefits from holding . . . insurance are delayed, the

⁷ See generally *id.*; see also Richard W. Cooper, *Spitzer's Allegations of the Anticompetitive Effects of Contingent Commissions: A Shot Truly Heard Around the World*, J. OF INS. REG. 83, 100 (2007).

⁸ See KENNETH S. ABRAHAM, INSURANCE LAW & REGULATION 56-57 (3d ed. 2000).

⁹ ERIC MILLS HOLMES, HOLMES' APPLEMAN ON INSURANCE 2D: LAW OF INSURANCE AGENTS § 47.5, at 326 (1998). See TOM BAKER, INSURANCE LAW AND POLICY 66 (2003); Colin Sammon, Comment, *Insurance Agent and Broker Liability: Crossing the Two Way Street*, 29 OHIO N.U. L. REV. 237, 238 (2002).

¹⁰ See Part II, *infra*.

probability of having a claim is hard to analyze, consumers do not get useful feedback about whether they are getting a good return on their insurance purchases, and mapping from what they are buying to what they are getting can be ambiguous.”¹¹ Consumers are therefore much more susceptible than commercial purchasers to being steered to insurance carriers they would not prefer under ideal market conditions.

Not only does such steering create mismatches between consumers and their insurers, but it undermines the competitiveness of consumer insurance markets as a whole. Although consumer insurance markets are ultra-competitive with respect to price,¹² they are remarkably non-competitive with respect to claims handling quality.¹³ Indeed, many consumer insurance markets appear to be characterized by insurer-side adverse selection, wherein price competition creates a race to the bottom among insurers with respect to claims handling quality.¹⁴ This Article argues that differential compensation contributes to this insurer-side adverse selection. By corrupting the objectivity of independent agents’ advice, differential compensation undermines the primary mechanism by which consumers can ordinarily overcome informational and cognitive limitations in assessing the quality of complicated financial products.

As such, this Article proposes that insurers who rely on independent agents to sell consumer lines of insurance should be prohibited from paying different rates of compensation to different agents for the sale of the same line of insurance. Such reform would be less radical than it may initially appear. Federal regulators have long regulated commissions

¹¹ CASS SUNSTEIN & RICHARD THALER, *NUDGE* 76-77 (2008).

¹² See J. DAVID CUMMINS, *DEREGULATING PROPERTY-CASUALTY INSURANCE* 2-3 (2002); Paul Joskow, *Cartels, Competition, and Regulation in the Property-Liability Insurance Industry*, RAND J. ECON. 375 (1973).

¹³ See Section II. B., *infra*.

¹⁴ Legal scholarship has long recognized that such price competition can cause firms to provide inefficiently poor quality when consumers cannot reliably evaluate quality due to information deficits or systematic cognitive limitations. See generally Oren Bar-Gill, *Seduction By Plastic*, 98 NW. L. REV. 1373, 1376 (2004); Jon D. Hanson & Douglas A. Kysar, *Taking Behavioralism Seriously: The Problem of Market Manipulation*, 74 N.Y.U. L. REV. 630 (1999); Jon D. Hanson & Douglas A. Kysar, *Taking Behavioralism Seriously: Some Evidence of Market Manipulation*, 112 HARV. L. REV. 1420 (1999).

for the sale of Medigap policies, and they recently announced their intention to do the same for Medicare Advantage programs.¹⁵ By extending these policies to the sale of all consumer insurance policies, lawmakers could provide consumers with the same protections that sophisticated commercial entities already enjoy. Even more importantly, they could enhance the competitiveness of consumer insurance markets as a whole.

I. INDEPENDENT INSURANCE AGENTS AND DIFFERENTIAL COMPENSATION

A. INSURANCE AGENTS IN CONSUMER INSURANCE MARKETS

Consumers can purchase insurance coverage directly from an insurer, or through either independent or captive agents. Captive agents are employees of a single insurer and only offer coverage with that carrier.¹⁶ By contrast, independent agents can write business with multiple insurers and consequently provide consumers with a choice of carriers.¹⁷ Such choice can be valuable for consumers, as insurers differ in terms of their reputations for claims handling, financial strength, risk management services, and scope of coverage offered.¹⁸ In addition to these variations in

¹⁵ See generally Omnibus Budget Reconciliation Act of 1990, Pub. L. No. 101-508, 104 Stat. 1388 (1990) (limiting agent compensation so that first year compensation may not be greater than twice renewal compensation, renewal compensation must be paid for at least 5 years, and replacement commissions may not be greater than renewal commissions for the product); Press Release, Center for Medicare & Medicaid Services, CMS Proposes New Protections for Medicare Beneficiaries in Medicare Advantage and prescription Programs (May 8, 2008) available at <http://www.cms.hhs.gov/> (describing proposed regulation that would “require Medicare Advantage organizations to establish commission structures for sales agents and brokers that are level across all years and across all [Medicare Advantage] plan product types”).

¹⁶ HOLMES, *supra* note 9, at 326.

¹⁷ Cummins & Doherty, *supra* note 4, at 375.

¹⁸ See *id.*; Schwarcz, *Beyond Disclosure*, *supra* note 1, at 296-97. Independent insurance agents market themselves primarily on the basis of their capacity to help consumers compare these variations in quality and pricing. As the website of their main trade organization explains, independent agents “work with you to identify the insurance . . . that [is] right for you... and use [their] access to multiple

“quality,” insurers employ differing underwriting criteria and strategies, resulting in price differentials even in highly price competitive marketplaces.¹⁹

Consumers may prefer independent agents over captive agents for other reasons as well. First, many insurers do not offer insurance directly to consumers or distribute their products through captive agents, meaning consumers who want to purchase policies from these insurers must go through an independent agent.²⁰ Second, because independent agents “own” their customer lists, insurers cannot directly solicit the agent’s clients or switch those clients to a different agent. Some have argued that this ownership gives independent agents a comparatively strong incentive to serve their clients, though empirical efforts have failed to confirm this theory.²¹

Of course, there are offsetting costs associated with purchasing coverage through an independent rather than captive agent. First, just as some insurers only provide coverage through independent agents, many popular insurers, such as State Farm and Allstate, only offer coverage through captive agents.²² In general, these insurers tend to be more

companies to deliver those products.” Independent Insurance Agents & Brokers of America – Consumer Information, <http://www.iaa.org>. A brochure designed by the National Association of Professional Insurance Agents similarly explains that “by shopping among various companies, your professional agent can find the best combination of coverage, price and service -- the best value for your insurance dollar.” National Association of Professional Insurance Agents, *Straight Talk about Choosing a Professional Insurance Agent*, <http://www.pianet.com/Publications/choosinganagentbrochure.htm>.

¹⁹ See, e.g., Meg Green, *Top of Their Game*, BEST’S REVIEW 26 (Dec. 2006) (describing how some of the most profitable property-casualty insurers focus on underwriting only particularly safe risks, and pass off some of the resulting cost savings to their insureds).

²⁰ See Laureen Regan & Sharon Tennyson, *Agent Discretion and the Choice of Insurance Marketing System*, 39 J. L. & ECON. 637, 639 (1996).

²¹ Helen Doerpinghaus, *An Analysis of Complaint Data in the Automobile Insurance Industry*, 58 J. Risk & Ins. 120 (1991); J. David Cummins & Stephen Weisbart, *The Impact of Consumer Services on Independent Insurance Agency Performance* (1977) (IMA Education and Research Foundation).

²² See Regan & Tennyson, *supra* note 20, at 638.

publicly visible, as large insurers can more easily support a captive distribution system and may also have greater advertising incentives.²³ Second, other things being equal, coverage purchased through captive agents will tend to be cheaper than coverage purchased through independent agents.²⁴ Because captive agents only work with one carrier, they spend less time on each sale, meaning that they receive lower commissions than independent agents.²⁵ These lower commission rates may result in lower premium rates for customers, as studies suggest that insurers pass through to consumers most of the cost of agent compensation.²⁶

B. DIFFERENTIAL COMPENSATION OF INDEPENDENT AGENTS

Independent agents are compensated through standard commissions on the premiums consumers pay for their coverage.²⁷ These “ordinary” commission rates have always varied based on the underlying line of insurance sold, as different lines of insurance require different levels of effort by insurance agents.²⁸ But, historically, these commission rates were relatively standard within specific insurance lines, as individual insurers offered a single commission rate to all agents. Although new insurers

²³ *Id.*

²⁴ Itzhak Venezia et. al., *Exclusive vs. Independent Agents: A Separating Equilibrium Approach*, 40 J. ECON. BEHAV. & ORG. 443, 444 (1999).

²⁵ See Regan & Tennyson, *supra* note 20, at 648-49.

²⁶ Cummins & Doherty, *supra* note 4, at 380-83. Competing factors, such as the improved quality of an insurer’s underwriting criteria which is caused by increased premiums, may offset this effect.

²⁷ See *id.* at 374. In property/casualty insurance markets, these commissions are generally the same each year that a consumer renews a policy, whereas commission rates tend to decrease over time for life insurance sales personnel. This creates its own conflicts of interest, which are beyond the scope of this Article.

²⁸ See *id.* at 374-75.

occasionally offered above-market rates to break into markets, competition ultimately ensured relatively uniform commissions within product lines.²⁹

In the last few years, the premium commissions that different insurers pay independent agents have begun to vary more significantly than in the past. Some insurers now negotiate their commission rates on an individual basis with agents, offering higher rates to agents that have historically directed a large volume of profitable business to the insurer.³⁰ As a result, many independent agents receive higher commission rates for selling policies from one insurer than another, despite competitive forces.

Even insurance agents who receive the same *premium commissions* from different insurers may nonetheless receive different *contingent commissions* from those insurers. Unlike differential premium commissions, insurers have long paid contingent commissions to independent agents.³¹ Contingent commissions are year-end bonuses that some insurers pay to independent agents based on the performance of the agent's book of business with that insurer.³² Most contingent commission contracts link this bonus to certain volume or profitability benchmarks for the agent's book of business. If the specified benchmarks are met, then the insurer pays the agent a contingent commission that usually is calculated based on the profitability and/or volume of the agent's book of business with that insurer.³³ In life and health markets, agents often receive these

²⁹ See Schwarcz, *Beyond Disclosure*, *supra* note 1, at 301.

³⁰ See, e.g., <http://www.chubb.com/marketing/chubb7450.html>; MetLife, *Supplemental Compensation Plan*, <http://www.whymetlife.com>; Rupal Parekh, *Hartford the Latest to Offer Supplementals as Replacements for Contingent Commissions*, *BUS. INS.*, 33 (July 2007).

³¹ See Fitzpatrick, *supra* note 1, at 3056 ("Contingent commissions have been used by insurers as an incentive mechanism for their agents for a century or more.").

³² *Id.*

³³ In general, the size of an intermediary's contingent commission is based on two variables: (1) the amount of insurance business that a particular intermediary refers to the insurer, as measured in total premiums; and (2) the profitability of that business, which is usually measured by the insurer's loss ratio on that business. In most cases, intermediaries are only entitled to contingent commissions if they meet threshold levels of both sales volume and profitability. See Jeffrey Wilder, *Competing for the Effort of a Common Agent: Contingency Fees in Commercial*

contingent commissions in the form of in-kind benefits, such as vacation trips, rather than monetary compensation.³⁴

However it is structured, differential compensation undermines independent agents' incentives to objectively present consumers with information about competing insurance options.³⁵ The reason is simple: they incentivize independent insurance agents to steer consumers to carriers based on considerations other than those customers' insurance needs and risk preferences. Most obviously, differential commissions encourage insurance agents to steer consumers to insurers who pay the highest commissions. But because differential commissions are almost always tied in some way to the volume and/or profitability of the agent's book of

Insurance 5 (U.S. Dep't of Justice, Antitrust Div., Econ. Analysis Group Working Paper No. EAG03-4, 2004), available at <http://ssrn.com/abstract=418061>. The loss ratio is the "ratio between premiums paid and losses incurred during a given period." BLACK'S LAW DICTIONARY 958 (7th ed. 1999). Premiums on both new policies and policy renewals are generally treated similarly in these calculations, which are almost always made on a yearly basis. Wilder, *supra*, at 5. In some cases, contingent commission arrangements may be based only on volume, not profitability. However, "the great majority of the arrangements covering the smaller intermediaries is based on the profitability of the business written or profitability and volume." Cummins & Doherty, *supra* note 4, at 379. Once intermediaries reach these qualifying levels, their commissions typically increase with better results along either dimension. See Wilder, *supra*, at 5.

³⁴ See, e.g., *Broker Compensation*, *supra* note 6, at 103-113 (testimony of F. James Ginnane) (describing various cruises to the Baltics, Sweden, Montreal and elsewhere that MassMutual paid based on annual production, and noting that "all of the carriers" he was familiar with offer similar trips), available at <http://www.ins.state.ny.us/agbrok/br-cmp-tran-buf.pdf>.

³⁵ Consumers who purchase insurance via a captive agent have already made a decision that they want to purchase their coverage with a particular carrier. This means they will often have already priced out several different carriers and, perhaps, asked neighbors or friends about their experiences with those carriers. By contrast, consumers who seek out coverage via an independent agent have typically not made any decisions about which carrier best suits their needs. Although they may have had a particular agent recommended to them, they generally do not even know which carriers the agent offers, much less the relative characteristics of those carriers. Rather, independent agents offer themselves to consumers as an alternative to comparison shopping among different insurers. They purport to do the comparative shopping for the consumer.

business with an insurer, they may also create more subtle steering incentives for agents.³⁶ For instance, they may lead agents to steer customers to an insurer that has a minimum-volume requirement on the cusp of being satisfied.³⁷ Alternatively, they may cause an agent who believes that a consumer is a "bad risk" to steer that consumer to an insurer with whom the agent does not have a differential commission arrangement tied to profitability.³⁸ Differential commissions may also increase premium costs for consumers.³⁹

II. THE DESIRABILITY OF A LEGAL RESPONSE TO DIFFERENTIAL COMPENSATION FOR INDEPENDENT AGENTS

Differential compensation of sales agents is common, and often understood to be relatively benign in many industries and market contexts. For instance, salespeople in retail stores may often receive special bonuses or in-kind benefits if they reach sales targets for particular products or brands. Like independent insurance agencies, such stores often carry multiple brands and consumers may rely on the advice of salespeople in making their decisions. Given that few suggest lawmakers regulate the compensation of sales personnel in these contexts, why would a different result be warranted in insurance markets?

Part of the answer is that consumer insurance markets are often regulated in ways that would be unthinkable in other markets. For instance, state insurance departments regulate product prices and designs and license salespeople and insurers.⁴⁰ Although the desirability of specific regulations

³⁶ See Schwarcz, *Beyond Disclosure*, *supra* note 1, at 297-301.

³⁷ See Wilder, *supra* note 33, at 19.

³⁸ Cummins & Doherty, *supra* note 4, at 386-89. For this reason, agents who steer "high-risk" consumers to certain insurers may theoretically undermine their client's interest by signaling to the insurer that particular consumers are relatively "high risk" and should thus be charged increased premiums. See Schwarcz, *Beyond Disclosure*, *supra* note 1, at 324-35.

³⁹ See Cummins & Doherty, *supra* note 4, at 383.

⁴⁰ See generally ETLINGER ET AL., STATE INSURANCE REGULATION 103 (1995).

is often contentious, the notion that insurance requires robust market conduct oversight is generally accepted. The reasons are two-fold. First, consumer insurance markets are uniquely susceptible to market failure for a variety of reasons, including the complexity of the underlying product, the cognitive limitations of consumers, the prevalence of information asymmetries, and various other external forces that distort the market by, for instance, mandating the purchase of coverage. Second, the consequences of such market failure are significant. Consumers who have inadequate coverage typically do not discover that fact until after they have suffered a loss, at which point they no longer have the ability to mitigate their damages.⁴¹

This Part applies these general rationales for insurance regulation to differential compensation arrangements in consumer insurance markets. It concludes the market forces that ordinarily limit the pernicious effects of differential commissions are unreliable in consumer insurance contexts. Similarly, it suggests that the consequences of the resulting market failure are significant, contributing to a race to the bottom over claims-handling practices in many consumer insurance lines.

A. DIFFERENTIAL COMPENSATION AND MARKET FAILURE

In ordinary product markets, an intermediary's temptation to push expensive or high-margin products is counter-balanced by the potential for market backlash.⁴² At least some consumers are likely to arrive at a store with some knowledge about competing product options, especially given the wealth of such information available on the internet. This is particularly true with big-ticket items – like high definition televisions or cars – about which consumers will often invest time in researching. Attempts to steer such consumers to inferior or overpriced products may backfire, resulting in those consumers shopping elsewhere and sharing their

⁴¹ Bob Works, *Excusing Nonoccurrence of Insurance Policy Conditions in Order to Avoid Disproportionate Forfeiture*, 5 CONN. INS. L. J. 505, 583-84 (1998).

⁴² Howell Jackson, *The Trilateral Dilemma in Financial Regulation*, in IMPROVING THE EFFECTIVENESS OF FINANCIAL EDUCATION AND SAVINGS PROGRAMS (Lusardi, Annamaria ed., forthcoming 2009); Howell E. Jackson & Laurie Burlingame, *Kickbacks or Compensation: The Case of Yield Spread Premiums*, 12 STAN J.L. BUS. & FIN. 289 (2007).

negative impressions with friends and family.⁴³ Although sales personnel may attempt to target uninformed consumers, such an approach can be risky as it may be hard to distinguish between informed and uninformed consumers. And even consumers that do end up purchasing inferior or over-priced products will often fail to discover this in the course of using their product.⁴⁴ Such consumers will not only hesitate before returning to the store, but they too may talk to family and friends about their negative experience.

To be sure, these market forces hardly eliminate sales contests and inducements that lead to slanted advice – there will always remain sleazy car salesmen, stores that sell over-priced and useless gadgets, and chains that push consumers to purchase over-priced accessories that add little to the overarching product. But the prospect that routine government intervention in these contexts could efficiently improve matters is slim. As this Section shows, these market forces that ordinarily protect consumers from excessive steering work poorly in consumer insurance markets.

i. *Information in Consumer Insurance Markets*

Unlike consumers in most markets, insurance consumers have access to few, if any, accurate measures of an insurer's reliability in paying claims fairly and efficiently.⁴⁵ It is, for instance, impossible for consumers to find out how often individual insurers pay claims within 30, 60, 90, or 120 days of a claim being reported; how frequently they deny claims; how frequently they are sued for payment or found guilty of bad faith; and how

⁴³ For a general discussion of the role of reputation in disciplining sellers' behavior, see Benjamin Klein & Keith Leffler, *The Role of Market Forces in Assuring Contractual Performance*, 89 J. POL. ECON. 615, 616 (1981).

⁴⁴ In economic parlance, insurance policies are thus "credence goods" because most consumers cannot evaluate their quality even after they purchase the policy. See Richard Craswell, *Interpreting Deceptive Advertising*, 65 B.U. L. REV. 657, 720-21 (1985) (explaining the differences between search goods, experience goods, and credence goods in economic and legal literature).

⁴⁵ "Information about the reliability of different insurers is hard to come by [and] the quality of insurance coverage is almost impossible to assess without an expert." KENNETH ABRAHAM, *DISTRIBUTING RISK: INSURANCE, LEGAL THEORY, AND PUBLIC POLICY* 176 (1986).

frequently policies are cancelled or non-renewed.⁴⁶ While consumers can look up how often complaints against specific insurers are lodged with state regulators, this data is notoriously unreliable and inconsistent.⁴⁷ Even the data published by *Consumer Reports* is highly limited, as it does not take into account the size and type of each consumer's claim and it is based on each consumer's subjective experience with the claims process.⁴⁸

Although consumers can, and do, carefully scrutinize premium differentials from different carriers, the significance of price differentials is almost impossible to assess without a corresponding understanding of the

⁴⁶ One recent proposal to the National Association of Insurance Commissioners (NAIC) would empower regulators to publicly disclose most of these data elements about the relative quality of insurers' claims handling, which they already collect. See PROPOSAL FOR CENTRALIZED DATA COLLECTION, NATIONAL ASSOCIATION OF INSURANCE COMMISSIONERS (Market Regulation Committee Proposal). Unfortunately, insurers have bitterly resisted the proposal under the guise of confidentiality and trade secrets. See Jim Connolly, *NAIC Insurer Conduct Data Scheme Riles Insurers*, NATIONAL UNDERWRITER, Sept. 25, 2008, available at <http://www.propertyandcasualtyinsurancenews.com/cms/nupc/Breaking%20News/2008/09/25-CONDUCTRULE-jc>; Chad Hemenway, *NCOIL Committee Votes Against NAIC Market Conduct Data Proposal*, BESTWIRE, Jul., 11 2008; Sean Carr, *NAIC Sets September Vote for Market Conduct Plan*, BESTWIRE, Jul. 28, 2008; Letter from Am. Health Ins. Plans, Am. Council of Life Insurers, Am. Ins. Ass'n, Blue Cross and Blue Shield Ass'n, Nat'l Ass'n of Mut. Ins. Cos., and Prop. Cas. Insurers Ass'n of Am. to Sandy Praeger, President of the Nat'l Ass'n of Ins. Comm'rs (May 27, 2008) available at http://www.naic.org/documents/committees_d_data_collection_comments_namic0527.pdf.

⁴⁷ See Daniel Schwarcz, *Redesigning Consumer Dispute Resolution: The American and British Approaches to Insurance Claims Conflict*, 83 TUL. L. REV. 735 (2009). Currently the Market Analysis and Priorities Working Group of the NAIC is working to develop better codes for the entire complaint data collection effort. See http://www.naic.org/committees_d_mapwg.htm (last visited February 8, 2009).

⁴⁸ Consumer Reports surveys thousands of consumers who filed claims and asks them to assess their satisfaction with the claims process. See *Consumer Reports Investigates, Surviving the Hard Market in Homeowners Insurance* Vol. 69, Issue 9, Consumer Reports. 36 (Sept. 2004); *Homeowners Insurance Report, The New Protection Game* Vol. 64, Issue 1, Consumer Reports, 16 (Jan. 1999).

differences in the underlying products.⁴⁹ A high-priced insurer may offer good coverage for a fair price, while a low-priced insurer may offer coverage that is poor, even relative to its seemingly low premiums. In fact, it is precisely for these reasons that independent agents choose to market themselves to consumers by focusing on their capacity to offer advice about competing carriers.

This lack of concrete information about the relative quality of different insurers undermines a key protection against aggressive steering in ordinary consumer markets. As described above, the fact that consumers ordinarily have the capacity to independently research and assess different product options limits the capacity of ordinary retail establishments to steer consumers to unfavorable deals. Moreover, it increases the prospect that attempting to do so will create market backlash, leading consumers who realize they are receiving poor advice to spread the word to others. But because most consumers simply do not have concrete information with which to assess the advice about the relative quality of carriers that insurance agents dispense, these protections are less robust in consumer insurance markets. This is particularly true given that insurance advice is hardly formulaic. The best insurance options for a customer may depend on numerous considerations, including the customer's risk tolerance, cash flow, preexisting relationships with carriers, and numerous other factors. This means self-serving advice can often be justified on some basis, and will rarely be obviously identifiable, even to experts.

Not only do consumers have a limited capacity to assess *ex ante* the quality of different carriers' coverage, but they also have a limited capacity to do so *ex post*. Unlike almost any other product, only a very small percentage of consumers end up using the insurance they purchase.⁵⁰ When they do, it is almost always for relatively small claims, even though the most important element of that insurance is the coverage it provides in cases of large losses.⁵¹ Finally, consumers that do submit claims to their insurers are typically ill-equipped to judge the extent to which their insurer

⁴⁹ Schwarcz, *Beyond Disclosure*, *supra* note 1, at 315.

⁵⁰ See Daniel Schwarcz, *A Products Liability Theory for the Judicial Regulation of Insurance Policies*, 48 WM. & MARY L. REV. 1389, 1413-15 (2007).

⁵¹ See *id.* at 1415.

lived up to its legal obligations.⁵² Consumers may therefore be susceptible to insurers' capacity to "tell a story" that appears to justify the refusal to pay a claim or a relatively low settlement of that claim.⁵³

This limitation in the capacity of consumers to assess insurance quality *ex post* further limits the power of market forces to address the steering that may result from differential compensation. Ordinarily, consumers who are successfully directed to inferior or overpriced products may discover this fact over time. Consumers who feel they were so victimized can not only choose to shop elsewhere in the future, but can talk to family and friends about their experience. Because most consumers who are steered to inferior insurance will never realize this fact, they will not exact these market penalties on agents who succumb to the temptation to maximize their compensation by directing consumers to inferior arrangements.

ii. *Consumer Decision-Making about Insurance*

In ordinary markets, consumers assess the desirability of different product options using a roughly rational process, at least in the aggregate. Especially when purchases involve big-ticket items, consumers are often willing to invest a significant amount of cognitive energy into making sure that they have thought through their options and selected a product that meets their needs and desires. As a result, salespeople can often exert only a minimal amount of pressure on shaping consumers' preferences. When salespeople push inferior or overpriced products, consumers may not only resist such practices, but may choose to avoid the establishment in the future and tell their friends and families of their experiences.

Two features of insurance markets substantially undermine this reasoned purchasing behavior, and the disciplining impact it has on agents' sales efforts. First, consumers typically purchase insurance as part of a larger event or transaction, such as taking a job, moving, or buying a home or automobile. Unlike with televisions, cars, or refrigerators, consumers do not typically decide that they can finally afford a new insurance policy, or

⁵² See Tom Baker, *Constructing the Insurance Relationship: Sales Stories, Claims Stories, and Insurance Contract Damages*, 72 TEX. L. REV. 1395, 1407-13 (1994).

⁵³ See *id.*

that their old policy is out of style, obsolete, or run down.⁵⁴ But the bundled decision-making that typifies such insurance purchases is both difficult and complicated, resulting in consumers “tend[ing] to adopt simpler choice strategies to cope with that complexity.”⁵⁵ Such simplistic strategies obviously enhance the capacity of sales agents to steer consumer decisions.

Second, empirical research has consistently demonstrated that consumers’ preferences concerning insurance are remarkably malleable.⁵⁶ Experimental research has established that framing effects can have important implications for consumers’ purchases of insurance policies. For instance, one study found that subjects were willing to pay more than twice as much for flight insurance covering “terrorism” and “mechanical failure” than they were willing to pay for flight insurance that would pay for losses for “any reason.”⁵⁷ Similarly, consumers tend to have bimodal responses to low-probability, high-cost risks, either dismissing them entirely or significantly overweighing their significance.⁵⁸ Which of these outcomes

⁵⁴ Evidence suggests that consumers rarely change carriers after they initially purchase a policy, especially outside of the auto insurance context. INSURANCE RESEARCH COUNCIL, PUBLIC ATTITUDE MONITOR 2001, Issue 2, at 5, fig. 2-3 (reporting that only 7% of homeowners or renters changed insurers in the last five years, but 23% of auto insurers did). When consumers do change insurers, they overwhelmingly cite price as the reason. *See id.* at 6, fig. 2-4.

⁵⁵ See Russel Korobkin, *Bounded Rationality, Standard Form Contracts, and Unconscionability*, 70 U. CHI. L. REV. 1203, 1223-34 (2003) (reviewing literature displaying this consumer tendency to choose simple choice strategies).

⁵⁶ Consumers’ decision-making processes about insurance are a complicated mix of intuitive, emotional, and rational responses that are susceptible to manipulation. *See* Horward Kunruether & Mark Pauly, *Insurance Decision-Making and Market Behavior*, 1 FOUNDATIONS AND TRENDS IN MICROECONOMICS 63 (April 2005); David Cutler & Richard Zeckhauser, *Extending the Theory to Meet the Practice in Insurance*, Brookings-Wharton Papers on Financial Services (2004); PAUL SLOVIC, *THE PERCEPTION OF RISK* 76-77 (2000).

⁵⁷ Eric J. Johnson, et al., *Framing, Probability Distortions, and Insurance Decisions*, 7 J. RISK & UNCERTAINTY 35, 39 (1993).

⁵⁸ Gary H. McClelland, et al., *Insurance for Low Probability Hazards: A Bimodal Response to Unlikely Events*, 7 J. RISK & UNCERTAINTY 95, 104, 108-09 (1993).

obtains often depends on the availability of the underlying risk. Thus, Californians' purchases of earthquake insurance generally increase significantly immediately after an earthquake occurs and then gradually decrease (until the next earthquake).⁵⁹ Finally, consumers' insurance decisions are significantly impacted by their affection for the item to be insured. In general, people prefer to insure against losses that involve high affect, even when holding constant the expected value of the insurance and the insured's level of wealth.⁶⁰ It is for precisely these reasons that insurers are among the heaviest advertisers of any industry.⁶¹

Given this malleability of consumers' insurance preferences, experienced or well-trained sales agents are likely to have a substantial capacity to steer consumers to insurers by helping to shape those consumers' preferences. This form of steering is unlikely to generate any market backlash, because it involves altering consumers' preferences. Often, this manipulation unambiguously impedes efficient market outcomes by skewing consumer assessments of objective information. This occurs, for instance, with the framing of a risk to increase a consumer's assessment of its likelihood. At the same time, other types of manipulation may admittedly operate on consumer insurance preferences in ways that are normatively ambiguous.⁶² Consider an agent who focuses on a consumer's affection for an item in order to increase her desire to insure against loss to that item. Evaluating the desirability of this result within a consequentialist

⁵⁹ See Paul Slovic et al., *Preference for Insuring Against Probable Small Losses: Insurance Implications*, 44 J. RISK & INS. 237, 249, 252, 254-55 (1977); HOWARD KUNREUTHER, *DISASTER INSURANCE PROTECTION: PUBLIC POLICY LESSONS* 26-27, 36-41 (1978).

⁶⁰ Christopher K. Hsee & Howard C. Kunreuther, *The Affection Effect in Insurance Decisions*, 20 J. RISK & UNCERTAINTY 141, 142-43, 148 (2000). Entire markets for insurance have flourished based on this principle: consider life insurance for children, which in most cases is irrational based on standard insurance theory.

⁶¹ See Baker, *supra* note 52, at 1404.

⁶² See Brett H. McDonnell, *Endogenous Preferences and Welfare Evaluations*, 5-6, 9-12 (Oct. 18, 2006) (unpublished manuscript), available at: <http://ssrn.com/abstract=933089>.

framework is difficult (if not impossible), because there is no exogenously-defined preference to serve as a benchmark for that evaluation.⁶³

iii. *Insurance Agents' Discrimination Between Sophisticated and Unsophisticated Purchasers*

In any consumer market, plenty of consumers will be relatively uninformed and therefore susceptible to inefficient steering. But these uninformed consumers are typically protected by their more informed counterparts. Because aggressive or misleading sales efforts that are directed at informed and engaged consumers can have negative effects on a business's reputation, uninformed or rationally ignorant consumers often benefit from the presence of their more informed counterparts when sales people cannot distinguish between the two.⁶⁴

Once again, though, this market protection against inefficient steering is less robust in insurance markets. Unlike most salespeople, insurance agents must discuss clients' personal situations in order to assess their coverage needs and facilitate insurer underwriting.⁶⁵ This process enhances agents' capacity to assess the relative sophistication of their consumers, and to offer advice accordingly. In fact, one of the earliest studies of contingent commission payments found just such a pattern of discrimination in a large independent insurance agency in Arizona: relatively engaged customers were less frequently directed to insurers that paid contingent commissions than customers who were less engaged with their insurance purchases.⁶⁶ Such consumer segmentation undermines one

⁶³ See *id.* at 12, 18.

⁶⁴ Alan Schwartz & Louis L. Wilde, *Intervening in Markets on the Basis of Imperfect Information: A Legal and Economic Analysis*, 127 U. PA. L. REV. 630, 636-38 (1979).

⁶⁵ Schwarcz, *Beyond Disclosure*, *supra* note 1, at 318.

⁶⁶ *Id.* at 317-18; see Wilder, *supra* note 33, at 2-3, 5, 7. The agency, which remained unidentified, employed eight agents with no ownership stake in the company and three "equity agents" who received a portion of the agency's profits. Because the contingent commissions that the company received were paid directly to the company, the three equity agents stood to gain more from maximizing contingent commissions than the non-equity agents. Additionally, only the equity agents handled "house" accounts, which (1) either originated in another agency that

of the core protections against undue steering in ordinary markets: the capacity of an informed minority to protect the interests of other consumers.⁶⁷

B. THE COST OF MARKET FAILURE FOR DIFFERENTIAL COMMISSIONS

Market failures, of course, are ubiquitous. And many of these market failures are better left alone than subjected to the expensive, and often ineffective (or worse), forces of government regulation. But that is not the case here. This Section argues that insurers' payments of differential compensation to independent agents facilitate a "race to the bottom" in consumer insurance markets through insurer-side adverse selection.⁶⁸ They do so by undermining the willingness of independent agents to inform consumers about insurers' claims handling practices or to counteract consumers' tendency to discount the value of quality claims handling.

the company subsequently acquired or were originally handled by an agent who retired, and (2) did not fit the portfolio or expertise of any non-equity agent. The defining characteristics of these house accounts strongly suggest that they were less sensitive than other agency customers to the level of service they received from their agent. This hypothesis was corroborated by the fact that house accounts were three times more likely than other accounts to pay their premiums directly to their insurer, rather than to pay them through the agency, indicating disengagement with their insurance agent. The study concluded contingent commissions significantly impacted the recommendations that the equity agents gave to their less responsive consumers, finding that "the prospect of contingency fees [led] equity agents to increase the frequency with which they place house accounts with insurers offering contingent commissions by more than 50%." *Id.*

⁶⁷ See Schwarcz, *supra* note 50, at 1406-08; R. Ted Cruz & Jeffrey J. Hinck, *Not My Brother's Keeper: The Inability of an Informed Minority to Correct for Imperfect Information*, 47 HASTINGS L. J. 635, 672, 674-75 (1996).

⁶⁸ BAKER, *supra* note 9, at 7. Just as insurer's lack of information about consumers can lead to adverse selection, consumers' lack of information about insurers can lead to the "insurer-side" adverse selection described above.

i. *Insurer-side Adverse Selection*

Part A described how insurance consumers' limited information on the relative quality of different insurance options and suspect decision-making about insurance can lead ostensibly independent agents to steer consumers to inferior insurers. But these two market conditions can also have the more general impact of undermining competition among insurers with respect to claims handling.

If a sufficiently large percentage of consumers are ill-informed about insurers' claims handling, insurers that pursue aggressive claims handling strategies (lemons) will profit more than other insurers. These insurers can pass on some of these profits to consumers in the form of lower premiums. In the long run, this will force other insurers to either drop out of consumer markets or, more likely, adopt low quality claims handling practices themselves.⁶⁹ By contrast, if a sizable number of consumers are cognizant of differences in insurers' claims handling, then some insurers will seek to appeal to these consumers by adopting a high price, high quality brand. That, in turn, could force other market players to compete over their own claims handling quality. Of course, insurers' quality/price mix would still vary, with different insurers appealing to consumers with different risk preferences. As a result, the market as a whole would compete along both of the two primary dimensions that define the insurance-policyholder relationship.

Through similar mechanisms, insurer-side adverse selection can occur if insurance consumers' decision-making causes them to under-value, or under-appreciate, differences in insurers' claims handling practices. There are strong reasons to suspect consumer decision-making about insurance generally has this character.⁷⁰ The relative value to consumers of high quality insurance depends on two considerations: (i) the likelihood they will suffer a potentially insurable loss, and (ii) the likelihood a low-quality insurer will poorly handle any such claim relative to a high-quality insurer. With respect to the former, research has consistently found that most people judge their own likelihood of suffering a loss to be lower than

⁶⁹ See generally Hanson & Kysar, *supra* note 14, at 630, 722, 724-25, 746-47 (exploring how consumers' under-estimation of risks can compel a similar race to the bottom with respect to those risks).

⁷⁰ Indeed, research has consistently found that there is a "systematic tendency for insurance in practice to differ from insurance in theory." Cutler & Zeckhauser, *supra* note 56, at 3.

the average such risk, so long as they retain even a minimal amount of control over the event.⁷¹ Thus, people in general are overly optimistic about their risk of being injured in an earthquake,⁷² being involved in a car accident,⁷³ suffering health problems,⁷⁴ and dying young.⁷⁵ For these reasons, they also generally buy less insurance against these risks than they should, especially when no outside force – such as legal mandates or loan terms – artificially increases demand.⁷⁶

Although less evidence exists as to how consumers evaluate the likelihood that a low-quality insurer will poorly handle a claim relative to a high-quality insurer, there are theoretical reasons to believe people will also tend to under-estimate this risk differential. In part, that is because consumers' choice of insurers involves precisely the minimal amount of control over an ultimate risk (the risk of a low-quality choice having negative consequences) that leads people in other contexts to believe their

⁷¹ A separate relevant strand of research has found that, when facing low-probability risks, people tend to either dismiss those risks entirely or overweigh the value of insurance against those risks. See SLOVIC, *supra* note 56, at 75, 77; McClelland, *supra* note 58, at 95, 108-109. This conclusion, however, has ambiguous implications for the extent to which consumers believe they will suffer an insurable loss, depending on the side of the bimodal distribution on which an insured risk falls.

⁷² Jerry M. Burger & Michele L. Palmer, *Changes in and Generalization of Unrealistic Optimism Following Experiences with Stressful Events: Reactions to the 1989 California Earthquake*, 18 PERSONALITY & SOC. PSYCHOL. BULL. 39, 40-1 (1992).

⁷³ David Dunning et al., *Ambiguity and Self Evaluation: The Role of Idiosyncratic Trait Definitions in Self-Serving Assessments of Ability*, in HEURISTICS AND BIASES 324 (Thomas Gilovich et al., eds., Cambridge Univ. Press 2002); KUNREUTHER, *supra* note 59, at 240.

⁷⁴ Neil D. Weinstein, *Unrealistic Optimism About Future Life Events*, 39 J. PERSONALITY & SOC. PSYCHOL. 806, 807 (1980).

⁷⁵ See Kyle D. Logue, *The Current Life Insurance Crisis: How the Law Should Respond*, 32 CUMB. L. REV. 1, 2, 4, 23 (2001-2002).

⁷⁶ See *id.* (noting the vast majority of Americans are under-insured against the risk of dying young); KUNREUTHER, *supra* note 59 (suggesting most Californians do not purchase earthquake insurance).

risk is lower than the average such risk.⁷⁷ Additionally, however, the actual difference between low and high quality insurers is ambiguous, in that it involves numerous considerations that are hard to definitively compare across insurers, even with all relevant information.⁷⁸ Research suggests people tend to interpret such ambiguous information in self-serving ways.⁷⁹ Given that high quality insurance unambiguously costs more than low quality insurance, this bias may theoretically manifest itself in consumers dismissing potential differences in claims handling quality.

Of course, the mere fact that economic conditions in insurance markets could theoretically lead to insurer-side adverse selection does not make it so. But many consumer insurance markets do appear to be characterized by some degree of insurer-side adverse selection, with few insurers pursuing high-quality, high-price strategies. Aside from the common (though often anecdotal) observations of commentators acknowledging this equilibrium,⁸⁰ significant evidence suggests prominent national insurers such as Unum/Provident, State Farm, and Allstate have each recently engaged in systematic, national efforts to cut claims payments

⁷⁷ See *supra* text accompanying notes 71-76.

⁷⁸ See *supra* text accompanying notes 45-53.

⁷⁹ Linda Babcock & George Loewenstein, *Explaining Bargaining Impasse: The Role of Self-Serving Biases*, 11 J. ECON. PERSP. 109, 111 (1997).

⁸⁰ See, e.g., *Pitman v. Blue Cross & Blue Shield of Okla.*, 217 F.3d 1291, 1296 (10th Cir. 2000) (observing Blue Cross had "a financial interest in denying claims in order to remain economically viable as well as competitive within the insurance industry"); BAKER, *supra* note 9, at 128 (collecting specific examples of seeming insurer opportunism); Schwarcz, *supra* note 50, at 1401-26; John Langbein, *Trust Law as Regulatory Law: The Unum/Provident Scandal and Judicial Review of Benefit Denials Under ERISA*, 101 NW. L. REV. 1315, 1331 (2007) ("Even when insurance is experience rated, the insurer still has an incentive to deny claims, because the market for insurance services is intensely competitive. Low-cost providers prevail over high-cost providers."). But see Alan O. Sykes, "Bad Faith" Breach of Contract by First Party Insurers, 25 J. LEG. STUD. 405, 418 (1996) (arguing that "any insurer who frequently refused to pay covered claims would likely soon develop a reputation for behaving in this fashion and lose customers," but acknowledging that "it is plausible that insurers might occasionally behave opportunistically without suffering a prohibitive reputational penalty").

to policyholders.⁸¹ Additionally, the insurance industry plays a significant role in limiting public access to information about different insurers' claims handling quality. For instance, insurers have collectively devoted immense energy and resources to ensuring that data about their claims handling quality, which is already collected by state insurance regulators, is not made publicly available.⁸² Similarly, studies of insurer marketing and advertising suggest that individual insurers do not publicly advertise any *concrete* information about the quality of their claims handling, preferring instead vague and unverifiable promises about trust (as well as concrete promises about price, of course).⁸³ In a market where insurers sought to compete over the quality of their claims-handling, one would expect that some insurers would prominently resist these trends.

ii. *The Role of Differential Compensation in Explaining Insurers' Race to the Bottom*

As described above, the two economic conditions that make insurer side adverse selection a plausible, and seemingly accurate, description of consumer insurance markets are (i) consumer ignorance about claims handling quality and (ii) under-appreciation of the significance of this variable. This Section suggests that differential compensation of independent insurance intermediaries is a key contributor to this equilibrium.

Consumer markets are ordinarily able to overcome informational problems through the evolution of a network of independent intermediaries that digest complicated data and objectively present consumers with advice.⁸⁴ This process allows consumers to make informed choices that

⁸¹ See JEFFREY STEMPEL, LITIGATION ROAD: THE STORY OF CAMPBELL V. STATE FARM (2008) (displaying State Farm's practices); Langbein, *supra* note 80, at 1318-21 (displaying Unum/Provident's practices); Consumer Watchdog, *Consumer Advocates Call for Refunds, Rate Reductions, For Allstate Policyholders After Company Releases Internal Documents Revealing Intentionally Underpaid Customers*, REUTERS, Apr. 7, 2008 (displaying Allstate's practices), available at <http://www.reuters.com/article/pressRelease/idUS187467+07-Apr-2008+PRN20080407>.

⁸² See documents cited *supra* note 46.

⁸³ See generally Baker, *supra* note 52.

⁸⁴ See Thomas F. Cotter, *Some Observations on the Law and Economics of Middlemen*, 1 MICH. ST. L. REV. 67, 69-70 (2006) (describing several economic

reflect their risk preferences despite their relative lack of understanding about the underlying market. Just as importantly, it improves the decision-making of less sophisticated consumers, by influencing insurers' reputations through word-of-mouth among consumers.⁸⁵

In the insurance context, objective and independent market intermediaries could accomplish these ends by digesting data on claims handling quality, along with repeated first-hand observation of insurers' practices, to accurately communicate information about insurers' claims handling practices. Such information gathering services are particularly significant in consumer insurance markets, not simply because of the dearth of public information on insurers' claims handling practices,⁸⁶ but also because few consumers could independently assess such information, even if it were publicly available. The quality of an insurer's claims handling is not a monolithic concept, and could be constructed in multiple ways, with differences in metrics appearing significant when they were not, or *vice versa*. For instance, data suggesting an insurer denied a relatively high percentage of claims, or a relatively high number of its consumers sue for coverage or complain to state regulators, might simply reflect the insurer's pool of policyholders, rather than its claims handling practices.⁸⁷

roles for middlemen, including assisting "consumers by reducing the cost of product search and evaluation, helping consumers to find the products that best fit their needs, and helping consumers to manage risk.").

⁸⁵ This role of market intermediaries in filtering and processing information for less sophisticated parties has been extensively discussed in debates on the efficient capital markets hypothesis, which is often imagined to achieve efficiency through a similar market intermediation mechanism. See generally Susanna Kim Ripken, *Predictions, Projections, and Precautions: Conveying Cautionary Warnings in Corporate Forward Looking Statements*, 2005 U. ILL. L. REV. 929 (2005) (discussing the role of market intermediaries in the efficient capital markets hypothesis).

⁸⁶ See *supra* Part II.A.

⁸⁷ Insurers have themselves seized on these difficulties in assessing claims handling data as one of their primary arguments against public disclosure. See Letter from Wiley Rein to Sandy Praeger, Pres. of the Nat'l Ass'n of Ins. Comm'rs 6 (April 16, 2008), available at http://www.naic.org/documents/committees_d_data_collection_comments_namic0416.pdf (resisting the public release of market conduct regulation, because "release of the information in raw form without the benefit of evaluation and interpretation would be unfair and potentially

Just as independent and objective sales agents can improve consumer information, they can also improve consumers' capacity to rationally and thoughtfully assess the trade-offs associated with purchasing relatively high quality insurance. By employing "debiasing" strategies, intermediaries may be able to counteract the tendency of consumers to under-appreciate the value of high quality coverage.⁸⁸ For instance, research suggests that people who are convinced that a potential loss is truly random generally no longer perceive they are relatively less likely than average to suffer from those losses.⁸⁹ By pointing out just how little control people have over the financial losses that are the subject of insurance, independent intermediaries could convince consumers to pay more for more reliable coverage. Similarly, independent agents might be able to concretize information about insurers' relative claims handling, thereby limiting the ambiguity of risk differentials that can trigger a self-serving interpretation of information. Even if independent intermediaries could not neutralize these biases, they might be able to counteract them.⁹⁰ For instance, independent agents could attempt to enhance consumers' evaluations of the risks attendant to low quality coverage by vividly describing these risks. Increasing the availability of risks can counteract consumers' tendency to underestimate them.⁹¹

Differential compensation undermines these market intermediation mechanisms by distorting the objectivity of the advice independent

damaging to insurers, and misleading to policyholders, investors, and the public at large.").

⁸⁸ See Christine Jolls & Cass R. Sunstein, *Debiasing Through Law*, 35 J. LEGAL STUD. 199, 204-05 (2006).

⁸⁹ Colin Camerer & Dan Lovallo, *Overconfidence and Excess Entry: An Experimental Approach*, 89 AM. ECON. REV. 306, 307 (1999); David Dunning, Chip Heath & Jerry M. Suls, *Flawed Self-Assessments: Implications for Health, Education, and the Workplace*, 5 PSYCH. SCI. PUB. INTEREST 69, 80 (2004) ("One of the strongest moderators of unrealistic optimism is perceived control. The greater a person's perceived control over an event or its outcome, the stronger the person's optimistic bias.").

⁹⁰ See Johnson et al., *supra* note 57, at 48. ("[C]onsumers' decisions about insurance can be affected by distortions in their perceptions of risk and by alternative framing of premium and benefits.").

⁹¹ Jolls & Sunstein, *supra* note 88.

insurance agents offer to consumers.⁹² Agents' capacity to mitigate consumer ignorance about insurance and debias consumers in ways that promote thoughtful consideration of insurance quality depends on agents prioritizing the interests of those clients. When intermediaries are incentivized to steer consumers to insurers in order to maximize their compensation, they are also encouraged to manipulate consumer preferences and impressions to achieve this outcome. This short-circuits the ordinary market solutions to informational and decision making problems in complex consumer markets. As a result, even insurers that are interested in cultivating a high-price, high-quality market strategy have limited vehicles for effectively communicating this strategy to potentially interested consumers. This creates circumstances under which insurer-side adverse selection with respect to claims handling can (and seemingly does) flourish.

III. CRAFTING AN EFFECTIVE RESPONSE TO DIFFERENTIAL COMPENSATION

Given the need for reform described in Part II, this Part briefly concludes by considering a simple legal intervention in insurance markets that resembles measures adopted in the federally-regulated markets that relate to Medicare.⁹³ That reform would limit insurer compensation of independent agents selling consumer lines of coverage to premium-based commissions, and would require insurers to pay a single, flat commission rate to all independent agents in their distribution networks. It would not mandate any particular commission rate, allowing insurers to choose the rate they wanted to offer to their independent agents. Insurers could set different premium commission rates for different lines of insurance, reflecting the fact that different product lines require different levels of effort for agents. Additionally, insurers could pay different commission rates to independent agents in different states to account for premium and cost of living differences across states.

Such reform would largely eliminate the distorting potential of differential compensation, leaving independent intermediaries without significant financial reasons to promote the policies of one insurer over another. Although some insurers might offer slightly higher commission rates than others, competition would ensure that these differentials would

⁹² See Section II.A, *supra*.

⁹³ See Press Release, *supra* note 15.

generally be quite small.⁹⁴ To the extent that differentials in commission rates persisted, they would be much less problematic than current commission differentials. Rather than rewarding individual preferred agents who steered consumers to a particular insurer, they would reflect an insurer's decision to offer above market commissions to all independent agents who sell a particular product line. As noted earlier, such a strategy might be justifiable for new entrants in a market seeking to establish a customer base.⁹⁵ Moreover, a high commission strategy might also be sensible for insurers offering high-price, high-quality products if the sale of such products requires comparatively more effort. Indeed, some economics literature suggests that sales agents in the consumer electronics industry may receive higher commission rates, on a per-dollar basis, for the sale of high quality products than low quality products for this reason.⁹⁶

Not only would a flat compensation rate for an insurer's independent agents help to solve the problems identified in Part II, but it would do so while imposing few administrative costs. A ban on contingent commissions or other specific compensation arrangements, standing alone, only invites insurers to design compensation structures that retain the same basic incentivizing function, but technically comply with the ban. Insurers' switch from contingent commissions to "supplemental compensation" arrangements, which retain the same performance-based contingency structure, is illustrative.⁹⁷ Because of its simplicity, a mandatory flat rate of

⁹⁴ Tacit collusion among insurers in setting commission rates would be unlikely, given the number of insurers who rely on independent agents to distribute their products.

⁹⁵ See *supra* text accompanying note 29 (noting that this was one reason historically that insurers offered higher premium commission rates).

⁹⁶ See Ajay Kalra, Mengze Shi, Kannan Srinivasan, *Salesforce Compensation Scheme and Consumer Inference*, 655 *Management Science* (2003).

⁹⁷ See Sally Roberts, *Compensation Shake-Up Continues; Chubb Pays \$17M, Ends All Contingents*, *BUS. INS.*, Dec. 25, 2006 (noting that the Chubb agreement states that "a fixed commission paid to a producer, set prior to the sale of a particular insurance product, and that may be based on, among other things, the prior year's performance of the producer' is not considered contingent"). Although some have suggested that these newly-emerging arrangements avoid the conflicts of interest associated with contingent commissions because they are "retrospective rather than prospective," this argument is unpersuasive. The fact that supplemental compensation arrangements are retrospective merely shifts

compensation for all intermediaries avoids this inefficient gaming. Moreover, it would be easy to enforce because it would operate on insurers rather than intermediaries. There are obviously fewer insurers than intermediaries (making market conduct observation easier) and insurers are less likely to engage in outright fraud than individual intermediaries who have less to lose from doing so.

Of course, mandating that insurers pay their independent agents a single commission rate is significantly more intrusive than a disclosure-based response to the problem. Not only would it be more costly to employ than disclosure, but it might distort consumer insurance markets in ways that may be hard to measure, or even predict. Nonetheless, such an aggressive intervention is prudent.

First, merely enhancing the disclosure requirements of independent agents is unlikely to mitigate the risk of steering, and the attendant risks of insurer-side adverse selection. Although I develop the limits of a disclosure-based regulatory response elsewhere,⁹⁸ the basic argument is simple: as described above, the reason that market forces do not prevent inefficient steering is that consumers generally have a limited ability to independently assess their insurance options. Merely informing consumers that their intermediaries may have a conflict of interest does nothing to address this fact. Of course, such disclosure could facilitate an agent's capacity to eschew differential compensation as a marketing technique.⁹⁹ But such efforts would be unlikely to prove profitable because consumers would have little sense of the value of such neutrality.

It is for precisely these reasons that compensation practices in consumer insurance markets have not, in fact, changed since 2004, despite the very public revelation of agents' conflicts of interest at that time and the adoption of mandatory disclosure laws in a number of states since.¹⁰⁰ This is particularly noteworthy given that numerous intermediaries in commercial insurance markets have voluntarily disclaimed differential

forward the potential pay-off to intermediaries of steering customers to sub-optimal insurance. See Schwarcz, *Beyond Disclosure*, *supra* note 1, at 292.

⁹⁸ Schwarcz, *Beyond Disclosure*, *supra* note 1.

⁹⁹ As noted above, this is precisely what has happened in commercial insurance markets.

¹⁰⁰ See Fitzpatrick, *supra* note 1, at 3064; Cooper, *supra* note 8, at 100.

compensation to recruit and retain new clients.¹⁰¹ Simply put, if potential clients are not attuned to the importance of unbiased advice or the ways in which advice can be distorted by incentive structures, they will not be swayed to change their behavior by competitors' promises of neutrality.

In fact, regulations of structurally similar conflicts of interest in other industries have often gone beyond disclosure-based strategies for precisely these reasons. Differential compensation of insurance intermediaries is one form of a common type of regulatory problem, coined a "trilateral dilemma."¹⁰² In a trilateral dilemma, an end-service provider compensates a market intermediary in order to induce the intermediary to steer consumers' business to the end-service provider.¹⁰³ Regulations of such side payments often do more than merely require disclosure, for the precise reasons developed above. Examples include prohibitions against certain side payments to real estate settlement providers,¹⁰⁴ limitations on side payments that brokerage firms can pay to investment managers,¹⁰⁵ and limitations on attorneys' receipts of side-payments for referrals to other attorneys.¹⁰⁶

Second, none of the proposed economic rationales for differential compensation appreciably enhance the efficiency of consumer insurance markets.¹⁰⁷ The most significant such potential benefit of differential

¹⁰¹ See *supra* text accompanying note 3.

¹⁰² See Schwarcz, *Beyond Disclosure*, *supra* note 1, at 312-19; Jackson, *supra* note 42.

¹⁰³ Jackson, *supra* note 42..

¹⁰⁴ See 12 U.S.C. § 2607 (2006).

¹⁰⁵ See D. Bruce Johnsen, *Property Rights to Investment Research: The Agency Costs to Soft Dollar Brokerage*, 11 YALE J. ON REG. 75, 82-83 (1994); see also 15 U.S.C. § 78b (2008).

¹⁰⁶ See Model Rules of Prof'l Conduct R. 7.2(b) (2009).

¹⁰⁷ Aside from the enhanced underwriting theory addressed in the text, contingent commissions have also been defended because they: (i) may expand coverage for non-verifiable losses, Neil A. Doherty & Alexander Muermann, *Insuring the Uninsurable: Brokers and Incomplete Insurance Contracts* 18 (Ctr. for Fin. Studies, Working Paper Nov. 24, 2005) available at http://www.ifk-cfs.de/papers/05_24.pdf; (ii) protect small agencies, Fitzpatrick, *supra* note 1, at 3042; and (iii) facilitate economies of scale by encouraging intermediaries to work

compensation is that it can improve the “front-line underwriting” of independent agents by giving them a stake in insurers’ profitability.¹⁰⁸ According to this theory, agents often possess information about the riskiness of customers that insurers cannot directly observe, as they interact directly with their customers and may have long standing relationships with them. Differential compensation that is linked to insurer profitability gives agents an economic reason to convey truthful information to the insurer. Alternatively, such compensation may facilitate improved underwriting simply by causing an agent who believes that a consumer is a “bad risk” to steer that consumer to a different insurer that does not pay differential commissions.

Whatever purchase this theory may have in commercial insurance markets, it is simply implausible in the context of consumer insurance lines. The theory assumes agents do indeed have important underwriting information about their clients that insurers cannot observe directly. But insurer underwriting in consumer insurance markets is generally standardized and based on simple and easily administrable algorithms.¹⁰⁹ Even if independent agents did possess information that could not be captured in an insurance application, it is unlikely that insurers would find

with fewer insurers, Cummins & Doherty, *supra*, note 4, at 386-89. For reasons developed in Schwarcz, *Beyond Disclosure*, *supra* note 1, at 305-11, these justifications are not persuasive. A final defense of contingent commissions – that they help small insurers to break into the market – is not in conflict with the proposal suggested herein, which would permit insurers to offer above-market premium commission rates.

¹⁰⁸ Cummins & Doherty, *supra* note 4, at 386-89; *see also* Regan & Tennyson, *supra* note 20, at 639 (“The agent is the first contact the insurer has with a potential policyholder and may be able to obtain information about the customer which would be difficult or costly for the firm to verify. It is widely acknowledged that agents often employ subjective criteria in evaluating insurance applicants.”).

¹⁰⁹ *See* ABRAHAM, *DISTRIBUTING RISK*, *supra* note 45, at 78 (“[A]n efficient classification system does not strive to make its premiums equal expected costs beyond the point where that goal is worth achieving.”). RICHARD V. ERICKSON, AARON DOYLE & DEAN BARRY, *INSURANCE AS GOVERNANCE* 241 (2003) (“Individual companies are increasingly less likely to undertake their own home inspection or direct field investigations of an applicant. Instead, more risk assessment is centralizing into data system operated by information service companies that supply the insurance industry.”).

incorporating that information into their underwriting to be cost efficient.¹¹⁰ This is especially true given the lack of adverse selection in most consumer insurance markets.¹¹¹

Of course insurance markets, like all markets, change over time. Thus, rationales for differential compensation that may not be compelling now may prove significant later. Consequently, any market intervention should be accompanied with continued monitoring and supervision. But the need for continuous re-assessment does not absolve lawmakers from ignoring conflicts of interest in consumer insurance markets that have been addressed in commercial and federally-regulated insurance markets. The failure of state lawmakers to act not only undermines the efficiency of consumer insurance markets, but it blunts the claim that consumer protection is best secured through the continuation of state-based insurance regulation.¹¹²

¹¹⁰ See H. LAURENCE ROSS, SETTLED OUT OF COURT: THE SOCIAL PROCESS OF INSURANCE CLAIMS ADJUSTMENTS 135 (1970) (explaining how insurers must adopt easily administrable rules of thumb to operate effectively).

¹¹¹ See Peter Siegelman, *Adverse Selection in Insurance Markets: An Exaggerated Threat*, 113 YALE L.J. 1223, 1224-5 (2004).

¹¹² See, e.g., Press Release, NAIC Still in Opposition to Federal Regulation (Jan. 20, 2009) (On file with the National Association of Insurance Commissioners).

The first question that arises is whether the contract is a contract of insurance. If it is, then the contract is subject to the provisions of the Connecticut Insurance Code. If it is not, then the contract is subject to the provisions of the Connecticut Contract Law. The Connecticut Insurance Code defines a contract of insurance as a contract in which one party agrees to indemnify another party against the risk of loss or damage to property or persons. The Connecticut Contract Law defines a contract as an agreement between two or more parties to do or not to do a certain thing. The Connecticut Insurance Code also defines a contract of insurance as a contract in which one party agrees to indemnify another party against the risk of loss or damage to property or persons. The Connecticut Contract Law also defines a contract as an agreement between two or more parties to do or not to do a certain thing. The Connecticut Insurance Code also defines a contract of insurance as a contract in which one party agrees to indemnify another party against the risk of loss or damage to property or persons. The Connecticut Contract Law also defines a contract as an agreement between two or more parties to do or not to do a certain thing.