31

Asset Sales in Good and Bad Times*

Marco Della Seta† Sebastian Gryglewicz‡

March 9, 2016

Abstract

This paper examines how preferences consistent with prospect theory affect asset sale decisions. We develop a continuous-time model in which a decision maker chooses when to sell an asset. The asset is subject to large risks in that infrequent shocks can substantially increase or decrease its value. The model predicts a disposition-type effect: the asset is sold later when payoffs are formulated in terms of losses (in bad times) than when they are formulated in terms of gains (in good times). Furthermore, the disposition-type effect becomes more pronounced when the variance of the payoffs is high. We confirm the model's predictions in laboratory experiments.

Keywords: Asset Sale, Exit, Risk, Prospect Theory.

JEL classification: G02, D81, C91.

---

*We also thank for feedback from seminar participants at University of Warwick, Tilburg University, Radboud University Nijmegen, and University of Lausanne, and participants of the ESA 2013 and EEA 2014.

†APG Asset Management. E-mail address: marco.dellaseta@apg-am.nl.

‡Erasmus University Rotterdam. E-mail address: gryglewicz@ese.eur.nl.

1 Introduction

On September 28, 2015, Shell announced the end of its oil discovery campaign in the Arctic after poor exploration results. The campaign lasted nine years and had an estimated cost $7bn. A few days later, two Italian entrepreneurs, Guido Martinetti and Frederico Grom, sold their ice-cream producing company, Grom, to the consumer goods giant Unilever. The two events, although different in scale, are closely related. They are opposite sides of one of the most important entrepreneurial and corporate decisions: exit. Shell abandoned an unsuccessful exploration project suffering substantial losses; Martinetti and Grom cashed in the results of several years of building a successful business. In both examples, a sale or exit offers a certain payoff against an alternative of a postponed risky payoff. In this paper, we develop a dynamic model of exit and test it in a series laboratory experiments, to study whether there is a systematic difference between exit in good times (i.e., when making gains) and exit in bad times (i.e., when suffering losses).

The behavioral significance of the distinction between gains and losses is a cornerstone of prospect theory (Kahneman and Tversky, 1979, 1992) and has been increasingly incorporated in economics and finance models. However, as Myagkov and Plott (1997) point out, the original formulation of prospect theory "addresses one-time decisions as opposed to repeated decisions or perhaps even substantially considered decisions that might take place in markets." There is surprisingly scant evidence on whether preferences consistent with prospect theory apply to relatively complex and dynamic economic problems. In this work, we attempt to partially fill this gap by studying whether and how the elements of prospect theory describe actual behavior in dynamic exit problems.

In the first step of our investigation, we develop a continuous-time model in which a decision maker has to decide whether and when to sell a risky asset. The asset is of unknown but constant quality, either high or low. If its quality is high, the asset produces a high payoff at a stochastic point in time, which provides an incentive to wait and postpone the decision to sell. However, waiting also carries a risk. Irrespective of its quality, the asset can be hit by a negative shock and, if this happens, the decision maker loses the opportunity to sell and obtains a lower payoff. The outside option to sell offers a constant price. If over a time interval the asset does not generate the high payoff, it becomes more evident that it is of the low quality. The decision maker's belief about the quality of the asset becomes less optimistic and the value of waiting erodes over time. Thus the problem produces a cutoff rule whereby the decision maker accepts the outside option after a certain period of waiting.

In this environment, we distinguish two settings, Good and Bad Times, depending on whether payoffs are formulated in terms of gains or losses. We solve the model in closed-form for the cases of preferences consistent with standard expected utility and of preferences consistent with prospect theory. In this second case, we assume that individuals display reference-dependent preferences, are loss averse, and are risk seeking for losses and risk averse for gains.[1] To make decisions in Good and Bad Times comparable, we pursue the following strategy. We calibrate the amount of initial wealth and the payoffs in such a way that the final amount of wealth in the three possible outcomes (the high payoff, the safe payment, or the low payoff) is the same in Good and Bad Times (that is, the initial level of wealth in Bad Times is appropriately higher compared to Good Times to compensate for the expected losses).

The model generates the following predictions. First, under standard expected utility preferences, the waiting time before selling the risky asset is the same in Good and Bad Times. The reason is that utility depends on the final level of wealth and we impose that this level is the same in Good and Bad Times. Then, the waiting time will be the same in both scenarios. Second, under prospect theory preferences, the safe payment is accepted

[1]In its original formulation, prospect theory (Kahneman and Tversky, 1979) also posits that individuals tend to overweight small probabilities and underweight high probabilities, the so-called probability weighting. Incorporating non-linear transformations of probabilities would make our dynamic problem untractable. As in other dynamic problems similar to ours (Kyle et al., 2006; Henderson, 2012) we omit this feature of prospect theory. We believe that this omission is not crucial for our investigation. First, our analysis focuses on qualitative predictions that are driven by the different risk attitudes in the gain and loss domains. Probability weighting would probably change the quantitative results without altering the qualitative conclusions. Second, in the experimental implementation we do not rely on extremely large or extremely small probabilities. Therefore, probability weighting is unlikely to play a major role in driving the results.

sooner in Good Times than in Bad Times. The explanation is intuitive. The decision maker is risk averse for gains but risk seeking for losses (the so-called reflection effect), and he is more willing to wait longer, and taking the risk to be hit by a negative shock, when payoffs are negative. We show that this behavior is not due to loss aversion and can also arise under piece linear utility due to reference dependence and time discounting. Third, a different risk attitude towards gains and losses implies that the distance between waiting times in the gain and loss domains is more pronounced when the variance of the payoffs is high.

Delaying the decision to sell when payoffs are negative is a behavior similar to the so-called disposition effect, that is the tendency of investors in the stock market to hold past losers and sell past winners (Shefrin and Statman, 1985).[2] For this reason, we refer to it as "disposition-type effect". However, our analysis is not explicitly directed towards stock trading. The asset's value in our model is not subject to frequent upward and downward movements, as a typical asset traded in the stock market, but is driven by large and infrequent shocks.[3] Our framework better describes different market settings in which asset values and prices are mostly moved by major news. One example is the real estate market in which property values may depend on the development of adjacent properties or on zoning changes which extend or limit potential usage. Another fitting example is oil or gas exploration companies whose value depends on finding exploitable deposits. The model can also describe a job search process, in which job offers arrive infrequently and individuals choose whether to accept, and stop the search, or to decline taking the risk of waiting for more attractive alternatives.

We test the model in a computerized laboratory experiment with student subjects. We run two treatments, one in the gain setting (Good Times) and one in the loss setting (Bad

[2]Barber and Odean (2011) state that the effect is a widespread phenomenon among individual investors and provide a review of existing evidence. The disposition effect has also been documented among institutional investors (Grinblatt and Keloharju, 2001) and in various markets [e.g., executive stock options in Heath, Huddart, and Lang (1999) and real estate in Genesove and Mayer (2001)]. The disposition effect has been documented in various markets [e.g., executive stock options in Heath, Huddart, and Lang (1999) and real estate in Genesove and Mayer (2001)].

[3]Such models are already studied in existing literature (Kyle, Yang, and Xiong, 2006; Barberis and Xiong, 2012; Henderson, 2012; Ingersoll and Jin, 2013)

Times). Each subject participates in only one treatment and plays multiple rounds. In each round, the variance of the payoffs can take three values, High, Medium, and Low, where High stands for high variance and so on. We test three hypotheses. Hypothesis 1 is about the existence of the disposition-type effect identified by the model and we posit that the risky asset should be sold earlier in Good Times than in Bad Times. In Hypothesis 2, we posit that in Good Times (Bad Times), the sale time should decrease (increase) with the variance of the payoffs. This is a consequence of the reflection effect in prospect theory preferences. Finally, Hypothesis 3 posits that observed disposition-type effect should be stronger when the payoff variance is high.

The experimental evidence is consistent with prospect theory preferences. In general, we confirm the disposition-type effect predicted by the model. Individuals in the Bad Times setting wait longer than in the Good Times setting. Also, within each setting, individuals wait longer when the payoff variance is low for gains, while they wait longer when the variance is high for losses. The disposition-type effect is stronger when the payoff variance is high, while it becomes insignificant when the variance is low.

Our paper contributes to the literature linking prospect theory with the disposition effect. This link began with the initial formulation of the disposition effect in a stock market setting (Shefrin and Statman, 1985). Although a number of studies consider prospect theory as a leading explanation for the disposition effect (e.g., Odean (1998), Frazzini (2006)), this relation is not yet clearly established. Several recent studies have attempted to formalize the intuitive link by formally modeling liquidation decisions under prospect theory. The existing theoretical contributions include Kyle, Yang, and Xiong (2006), Barberis and Xiong (2012), Henderson (2012), and Ingersoll and Jin (2013), who show that prospect theory preferences generate the disposition effect, but also Barberis and Xiong (2009), who show that prospect theory can for some parameters imply, in contrast, that investors are more inclined to sell past losers than past winners. Even if the theory can predict the disposition effect, it should be empirically tested as an explanation against alternatives. Existing tests using field data show mixed results (e.g., Odean, 1998; Kaustia, 2010; Ben-David and Hirshleifer, 2012). Ben-David and Hirshleifer (2012) point at specific factors that "severely confound" the disposition-effect tests based on happenstance data and at the tests' lack of ability to support or reject preference-based explanations (and other leading explanations). Controlled laboratory environments provide researchers unique opportunities to disentangle alternative theories. Our theoretical model reinforces the link between prospect theory and disposition effect and shows that a disposition-type behavior, a delayed decision to exit, can arise when the asset's value is subject to large risks. Furthermore, we provide experimental evidence in support of a prospect theory interpretation.

In the theoretical part, our paper is most closely related to the single-asset partial-equilibrium models of Kyle, Yang, and Xiong (2006) and Henderson (2012).[4,5] The main difference is that in our model, uncertainty is modeled with a Poisson process rather than a diffusion process. This means that the asset's value is driven by large, and relatively infrequent, shocks rather than by small and frequent movements. As either type of uncertainty may dominate for different assets, the two settings can be seen as complementary. Additionally, the Poisson setup makes our model simpler than the diffusion-based models.[6] This simplicity, hopefully, helps experiment participants to make informed decisions. Another difference is that our model is solved and our predictions hold for a general utility function. This generalization is important in our context as we do not need to assume a particular specification to interpret the experimental results. Finally, our model can be viewed as a dynamic version of some experiments in Kahneman and Tversky (1979) (as explained in Section 3). This again facilitates implementation in the laboratory and interpretation of the

[4] Barberis and Xiong (2012) and Ingersoll and Jin (2013) present models that allow for reinvestment; a similar comparison applies to these papers as to Kyle, Yang, and Xiong (2006) and Henderson (2012).

[5] More generally, sale or exit timing decisions have been studied in various applied settings. Examples of exit problems in good times are Pastor and Veronesi (2009), who investigate the decision of a private entrepreneur to sell the firm in an IPO, and Schwienbacher (2008), who examines venture capital exits through IPOs and trade sales in a model of innovation. Problems of exit in bad times in a dynamic stochastic environment are studied in Dixit (1990).

[6] Optimal stopping problems of a diffusion process become complex with non-smooth objective functions, as in the case prospect theory preferences. Both Kyle, Yang, and Xiong (2006) and Henderson (2012) are also technical contributions in solving the respective mathematical problems.

results.

In the experimental part, our paper is most closely related to Weber and Camerer (1998), who find evidence consistent with the disposition effect in a security trading experiment.[7] However, they suggests that their results could be partially driven by the fact that subjects might mistakenly believe that the value of the risky asset is mean-reverting. Since, as explained above, the asset in our model is subject only to a one-time jump in value and moves in one direction, we are able to rule out mean-reversion as a potential explanation of the results. More broadly, our work also adds to a group of papers that experimentally test optimal stopping models (Oprea, Anderson, and Friedman, 2009; Anderson, Friedman, and Oprea, 2011; Della Seta, Gryglewicz, and Kort, 2014).

2 The Model

2.1 Setup

We present a model in which a decision maker owns a risky asset and decides whether and when to sell it. Time is continuous and labeled by $t \in [0, \infty)$. The decision maker (DM henceforth) discounts future income at rate r and is endowed with an initial wealth w_0. Preferences are defined by a utility function strictly increasing in its argument. The asset is of unknown quality and is subject to large risks. The asset's quality is either high, when it yields a payoff H, or low, when it yields a payoff $L < H$. The initial probability p_0 that the quality is high is known to the DM.

The DM has to wait to obtain either of the payoffs, H or L. If the quality of the asset is high, there is a Poisson process with intensity μ that reveals its quality. Upon the arrival of this Poisson event, the high payoff is realized and the problem ends. Thus, in a time interval Δt, positive news about the quality of the asset arrives with probability $\mu \Delta t$. Additionally,

[7] Frydman, Barberis, Camerer, Bossaerts, and Rangel (2014) run Weber and Camerer's (1998) experiment and apply an alternative methodology by studying subjects' neural activity.

the asset value can be hit by an adverse shock after which it yields L irrespective of its quality. The arrival of a negative shock is modeled as an independent Poisson event with intensity ϕ. Random times at which payoffs H and L are realized are denoted by T^H and T^L, respectively.

In addition to waiting for risky payoffs from the asset, the DM has an opportunity to accept an outside option. This could be interpreted, for example, as selling the asset to an outsider. The payoff from the sale is M and we assume that $H > M > L$. The outside option M is available only before the payoffs H or L are realized. Thus, the risky asset potentially yields the high outcome H but also involves the risk of realizing L, which would be lower than the payoff from the outside option, M. At each point in time, the alternative for the DM is to choose between waiting and selling the risky asset. We interchangeably refer to the latter decision as a sale, liquidation, or exit.

Our model setup is similar to Miao and Wang (2007) and can represent a dynamic model of entrepreneurship with exit. More generally, the model is a variant of the bandit problem, which has found widespread application in economics and finance [e.g., studying innovation in Weitzman (1979), capital investment in Décamps and Mariotti (2004), or delegated innovation in Manso (2011)].

2.2 Learning

The DM has imperfect information about the asset's quality, but information is revealed over time. This generates a value of waiting. Here we describe the DM's learning process. The DM's posterior belief at time t that the asset is of high quality is denoted by p_t. In a time interval Δt, the asset can be revealed to be of high quality by generating a high payoff, H. If this does not happen, the DM updates his belief in a Bayesian fashion. According to Bayes' rule, after an interval Δt, the DM's posterior belief is:

$$p_t + \Delta p_t = \frac{(1 - \mu\Delta t)\, p_t}{(1 - \mu\Delta t)\, p_t + (1 - p_t)}.$$

Taking the limit for $\Delta t \to 0$ and rearranging yields the instantaneous change in belief:

$$dp_t = -\mu p_t (1 - p_t)\, dt. \tag{1}$$

Solving (1) yields:

$$p_t = \left(1 + \frac{1 - p_0}{p_0} e^{\mu t}\right)^{-1}. \tag{2}$$

The learning process has the following properties. First, p_t decreases deterministically over time. If the asset does not yield a high payoff, the DM becomes progressively less confident that the quality of the asset is high. Second, the speed of learning increases with μ. When μ is larger, the fact that in a certain time interval the asset is not revealed to be of high quality is a more reliable signal and the DM updates his belief more rapidly. Third, the posterior belief p_t converges to zero in the limit $t \to \infty$. If a positive signal about the asset's quality does not arrive, the DM eventually becomes almost certain that the payoff from investment is L.

2.3 The Decision Problem

In this subsection, we solve the DM's problem. As a benchmark, we assume that the DM has standard expected utility preferences. Specifically, the DM derives utility from his final level of wealth, denoted by w, and the utility function $u(w)$ is differentiable, strictly increasing and concave in the relevant domain, $u'(w) > 0$ and $u''(w) < 0$. Depending on the realization of the payoff, L, M, or H, the DM's final wealth is equal to either $w_L = w_0 + L$, $w_M = w_0 + M$, or $w_H = w_0 + H$. Clearly, it holds that $u(w_H) > u(w_M) > u(w_L)$.

The DM must decide when to sell the asset. Intuitively, by waiting, the DM receives information about the quality of the asset. If time proceeds and the asset does not yield a high payoff, the probability that the asset is of high quality becomes eventually so low that it is optimal for the DM to accept the outside option and receive w_M. Formally, the DM's problem is to find a stopping time τ that maximizes the following objective:

$$\max_{\tau} E\left[u(w_M)e^{-r\tau}1_{\tau<T^H\wedge T^L} + u(w_H)e^{-rT^H}1_{T^H<\tau\wedge T^L} + u(w_L)e^{-rT^L}1_{T^L<\tau\wedge T^H}\right]. \quad (3)$$

The term in the brackets is simply the discounted value of the payoff that realizes first ($x \wedge y$ denotes the minimum of x and y). Before any payoff, the DM's belief, p, is the only state variable. This suggests that solving the problem entails finding the posterior belief p^* such that, if $p \leq p^*$, it is optimal for the DM to accept the outside option while, if $p > p^*$, it is optimal to further wait for the high payoff. Since Equation (2) establishes a one-to-one correspondence between belief and time, the solution of the problem can be equivalently expressed in terms of calendar time. The time at which it is optimal to accept the outside option will be denoted by τ^*.

Let $V(p)$ denote the DM's value function in the continuation region, that is, for $t < (\tau \wedge T^H \wedge T^L)$. V is a function of the belief process with the dynamics described by (1). Standard arguments imply that $V(p)$ satisfies the following Bellman equation:

$$rV(p) = -\mu p\,(1-p)\,V'(p) + \mu p\,(u(w_H) - V(p)) + \phi\,(u(w_L) - V(p)). \quad (4)$$

Equation (4) is interpreted as follows. The DM is willing to postpone the acceptance of the outside option if the normal return on his value function $rV(p)$ (the left-hand side) equals the expected change of the value function (the right-hand side). The latter term consists of three components. The first, $-\mu p\,(1-p)\,V'(p)$, measures the effect of the instantaneous change in belief. The second, $\mu p\,(u(w_H) - V(p))$, accounts for the arrival of the high payoff in case the asset is of good quality, where μp is the conditional instantaneous probability of this happening, while $u(w_H) - V(p) > 0$ is the net payoff. The third component, $\phi\,(u(w_L) - V(p)) < 0$, accounts for the effect of the arrival of a negative shock, which leaves the DM with a low payoff.

Equation (4) must be solved under the following boundary conditions:

$$V(p^*) = u(w_M), \tag{5}$$

$$V'(p^*) = 0. \tag{6}$$

Equation (5) is the value-matching condition, which simply requires that the DM's value equals the utility deriving from the outside option at a threshold p^*. Equation (6) is the so-called smooth-pasting condition. It is standard in optimal stopping problems and ensures that the threshold p^* is optimally chosen (see Dumas, 1991).

Plugging (5) and (6) in (4) yields:

$$ru(w_M) = \mu p^* \left(u(w_H) - u(w_M)\right) + \phi \left(u(w_L) - u(w_M)\right).$$

Solving for p^*, we obtain the threshold that induces the DM to sell the risky asset:

$$p^* = \min\left[p_0, P\right], \tag{7}$$

where

$$P = \frac{(r+\phi)\, u(w_M) - \phi u(w_L)}{\mu \left(u(w_H) - u(w_M)\right)} > 0. \tag{8}$$

The optimal timing τ^* can be found by plugging p^* to (2) and solving for t:

$$\tau^* = \frac{1}{\mu} \log\left[\frac{p_0 \left(1 - p^*\right)}{p^* \left(1 - p_0\right)}\right]. \tag{9}$$

So far, we have assumed that the DM makes the exit decision on the basis of the final level of wealth associated with the different outcomes. Experimental evidence on static lotteries suggests a different preference model. According to prospect theory, individuals' risk attitudes depend on whether the decision problem is formulated in terms of gains or losses with respect to a reference point, typically the initial wealth (Kahneman and Tversky, 1979).

Consistently with prospect theory preferences, we assume now that the DM values payoffs with a value function $v(x)$ that depends on final gains or losses x relative to the reference point of initial wealth, w_0. Using x as the argument in the value function is consistent with the realization utility model of Shefrin and Statman (1985) and Barberis and Xiong (2012) who posit that DMs derive utility from *realized*, in our case, final, losses or gains. Besides reference dependence, prospect theory preferences exhibit risk aversion in gains and risk seeking in losses. This means that $v(x)$ is concave in gains ($v''(x) < 0$ if $x \geq 0$) and convex in losses ($v''(x) > 0$ if $x < 0$), a property sometimes referred to as diminishing sensitivity or reflection effect. Furthermore, prospect theory suggests a larger sensitivity to losses than to gains of the same magnitude, that is, loss aversion. This property means that the utility function in the loss domain is steeper than in the gain domain.

To study the consequences of prospect theory on asset sale decisions, we characterize payoffs separately in the two domains. In the gain domain, referred to as Good Times, payoffs for the three possible outcomes are $\overline{H}$, $\overline{M}$, and $\overline{L}$, respectively, where $\overline{H} > \overline{M} > \overline{L} > 0$. The initial wealth is $\overline{w}_0$. In the loss domain, referred to as Bad Times, payoffs are $\underline{H}$, $\underline{M}$, and $\underline{L}$, where $0 > \underline{H} > \underline{M} > \underline{L}$, while the initial wealth is $\underline{w}_0$. Let p_G^* and p_L^* denote the thresholds in Good and Bad Times, and τ_G^* and τ_L^* denote the corresponding optimal stopping times to sell the risky asset. Following the same steps as in the previous section, the optimal thresholds can be found as

$$p_G^* = \min\left[p_0, P_G\right] \text{ and } p_L^* = \min\left[p_0, P_L\right] \tag{10}$$

and, using (9),

$$\tau_G^* = \frac{1}{\mu}\log\left[\frac{p_0\left(1-p_G^*\right)}{p_G^*\left(1-p_0\right)}\right] \text{ and } \tau_L^* = \frac{1}{\mu}\log\left[\frac{p_0\left(1-p_L^*\right)}{p_L^*\left(1-p_0\right)}\right], \tag{11}$$

where

$$P_G = \frac{(r+\phi)\, v(\overline{M}) - \phi v(\overline{L})}{\mu\left(v(\overline{H}) - v(\overline{M})\right)} > 0, \tag{12}$$

$$P_L = \frac{(r+\phi)\, v(\underline{M}) - \phi v(\underline{L})}{\mu\left(v(\underline{H}) - v(\underline{M})\right)} > 0. \tag{13}$$

3 Model Implications

3.1 Gains, Losses, and Asset Sale Decisions

To evaluate the implications of the different risk attitudes on the sale decision, we adopt the following strategy. We assume that despite differences in the payoff levels, the final level of wealth does not differ in the gain and loss domains. More formally, we make the following assumption.

Assumption 1 The following equalities are satisfied:

$$\overline{w}_0 + \overline{H} = \underline{w}_0 + \underline{H},$$

$$\overline{w}_0 + \overline{M} = \underline{w}_0 + \underline{M},$$

$$\overline{w}_0 + \overline{L} = \underline{w}_0 + \underline{L}.$$

Assumption 1 requires that $\underline{w}_0 - \overline{w}_0 = \overline{H} - \underline{H} = \overline{M} - \underline{M} = \overline{L} - \underline{L} > 0$ and implies that the final wealth for the three possible outcomes is the same in both Good and Bad Times. The exit problems in the gain and loss domains can be viewed as dynamic versions of Problems 11 and 12 in Kahneman and Tversky (1979).

As a benchmark, let us consider first an implication of Assumption 1 when the DM has a utility function that satisfies the axioms of the expected utility theory.

Proposition 1 *If the utility function satisfies the axioms of the expected utility theory, then the optimal timing of asset sale is the same in the gain and loss domains, i.e.,* $\tau_L^* = \tau_G^*$.

Proof. See Appendix A.1. ∎

Proposition 1 says that in the standard expected utility framework the optimal exit time is the same for Good and Bad Times. The intuition is straightforward. Assumption 1 makes the final wealth equal across the gain and loss domains, so when the DM bases the decisions on the final wealth, the timing is the same irrespective of the domain.

Let us now consider what happens if the DM has a value function that features reference dependence with respect to initial wealth and is concave in the gain domain and convex in the loss domain.

Proposition 2 *Under prospect theory preferences, the risky asset is sold later in Bad Times than in Good Times, i.e.,* $\tau_L^* \geq \tau_G^*$.

Proof. See Appendix A.2. ∎

Proposition 2 implies that prospect theory preferences make the DM more reluctant to realize losses. Therefore, the risky asset is sold later in Bad than in Good Times. This behavior is consistent with the disposition effect, and we refer to it as disposition-type effect.

3.2 What Matters in Timing Decisions?

In this section we examine which elements of prospect theory may play a role in differential timing behavior in the loss and gain domains. Such understanding facilitates the interpretation of experimental results.

For the most part, we use the general formulation of prospect-theory preferences introduced in Section 2. However, to be able to derive more concrete predictions about the role of loss aversion, we consider two popular parametric specifications for the DM's value function.[8]

[8] The difficulty of making general statements with respect to loss aversion is related to the fact that several different measures of loss aversion have been suggested (see, e.g., Wakker and Tversky, 1993; Neilson, 2002; Köbberling and Wakker, 2005).

The first is the power function proposed by Tversky and Kahneman (1992):

$$v(x) = \begin{cases} x^{\alpha} \text{ if } x \geq 0 \\ -\lambda(-x)^{\beta} \text{ if } x < 0, \end{cases} \tag{14}$$

where $\alpha, \beta \in (0,1)$, and $\lambda > 1$ is a coefficient of loss aversion. The second is the two-piece exponential function used in Kyle et al. (2006):

$$v(x) = \begin{cases} \psi_1 \left(1 - e^{-\gamma_1 x}\right) \text{ if } x \geq 0 \\ \psi_2 \left(e^{\gamma_2 x} - 1\right) \text{ if } x < 0, \end{cases} \tag{15}$$

where $\psi_1, \gamma_1, \psi_2, \gamma_2 > 0$. The assumption $\psi_1\gamma_1 < \psi_2\gamma_2$ guarantees that the utility function is steeper in the loss domain around the reference point.

Combining different elements of prospect theory preferences, we have the following results.

Proposition 3 *The risky asset is sold later in Bad Times than in Good Times, i.e.,* $\tau_L^* \geq \tau_G^*$, *if the DM's preferences exhibit: (i) reference dependence, linear utility, and time discounting* $(r > 0)$*; or (ii) reference dependence and the reflection effect. Loss aversion in the forms exhibited in the value functions in (14) and (15) does not have an additional effect on timing.*

Proof. See Appendix A.3. ■

The first result of the proposition is that even under local risk-neutrality, the disposition effect arises due to reference dependence and time discounting. The intuition is simply that the DM prefers to realize losses later when they are discounted. This finding is consistent with a similar finding of Barberis and Xiong (2012) in a model in which the risky asset follows a diffusion process. However, since in our experiment the time horizon is short, discounting is unlikely to play any significant role. As the second part of Proposition 3 indicates, loss aversion also does not impact our disposition effect.[9] Thus we expect that the reflection

[9]The reason is that if all the payoffs, L, M, and H, are either negative or positive, then the loss aversion parameters cancel out from the optimal stopping problem.

effect is the feature that can manifest in different behavior in the loss and gain domains in the experiment.

Therefore, we explore further the implications of the reflection of the utility function around the reference point in the absence of discounting. If the DM is locally risk averse in the gain domain and risk seeking in the loss domain, decisions in the two domains can be affected differently by increased risk. We examine the influence of increased riskiness of the asset on exit timing. To isolate the risk effect, we define increased risk as increased spread between the H and L payoffs, keeping the safe payoff M and the ratio $(M-L)/(H-M)$ fixed. If $(M-L)/(H-M)$ is constant, then in absence of discounting, optimal timing decisions with risk-neutral preferences (with or without reference dependence) would not be affected by changes in such defined risk. We have the following result.

Proposition 4 *Suppose the DM exhibits prospect theory preferences and $r=0$. Then with increasing risk, the risky asset is sold sooner in Good Times, later in Bad Times, and the disposition-type effect is more pronounced.*

Proof. See Appendix A.4. ■

4 Experiment and Testable Hypotheses

The goal of our experimental study is to test the model's predictions. To do so, we designed a computerized laboratory experiment that recreates the exact setting of the model. Subjects played the roles of the model's decision maker endowed with some initial wealth and a risky asset. As in the model, the asset can generate a high or a low payoff at a random time in the future and an immediate safe payoff when the subject decides to sell the asset.

In the lab, we approximate continuous time by using a discretization with a time interval equal to 0.1 seconds. We set the probability of the risky asset being of high quality equal to 65% ($p_0 = 0.65$). The probability that the risky asset yields the high payoff if it is of high quality and the probability of a negative shock are both set equal to 0.0041 every 0.1

Treatment	Good Times				Bad Times			
	H	M	L	$\overline{w}_0$	$\underline{H}$	$\underline{M}$	$\underline{L}$	$\underline{w}_0$
High	30	10	0	5	0	−20	−30	35
Medium	24	10	3	5	−6	−20	−27	35
Low	16	10	7	5	−14	−20	−23	35

Table 1: Payoffs in the High, Medium, and Low payoff schemes for the Good and Bad Times treatments.

seconds, which corresponds to 4% every second. A 4% probability every second for the two exogenous events leaves the subjects sufficient room to wait before the game is terminated for external reasons.

We run two treatments, Good Times and Bad Times, in which payoffs are formulated in terms of gains and losses but the final wealth levels are the same (i.e., Assumption 1 is satisfied). Each subject participated in one treatment only. From the analysis in Section 3 and Proposition 2, we arrive at Hypothesis 1.

Hypothesis 1 The risky asset is sold earlier in the Good Times treatment than in the Bad Times treatment.

We let subjects play the game for 45 rounds. At each round, payoff values can change. Specifically, three different payoff schemes can occur with equal probability. In the three schemes, the ratio $(M-L)/(H-M)$ is constant but the variance of the payoff is different. The payoff schemes are called High, Medium, and Low depending on the level of the variance. Payoffs in the different schemes and treatments are summarized in Table 1. All payoffs are expressed in euros.

There are two reasons to let the payoffs change in every round. First, since the experiment lasts for 45 rounds, it helps to keep the attention of the subjects. Second, it allows us to investigate in more detail the predictive power of the model.

Following Proposition 4, we hypothesize the following differences in the observed behavior under Low, Medium, and High risk treatments.

Hypothesis 2 In the Good Times treatment, exit is soonest in High and latest in Low. In

	Predicted sale time, τ^*		
	High	**Medium**	**Low**
Good Times	3.90	9.15	12.88
Bad Times	22.65	19.22	16.89

Table 2: Theoretical predictions (in seconds) for the sale trigger for the Good and Bad Times treatments in the different payoff schemes using the utility parametric specification of Tversky and Kahnemann (1992).

the Bad Times treatment, exit is soonest in Low and latest in High.

Hypothesis 3 The strength of the disposition-type effect increases with the variance of the payoffs (from Low to High).

Hypotheses 2 can be tested *within* subjects and is a consistency check for the decision model with the reflection effect around the reference point. If individuals are risk averse in the gain domain, then the timing of the sale decision should be negatively related to the variance of payoffs in the Good Times treatment. Symmetrically, if individuals are risk seeking in the loss domain, the timing of the sale decision should be positively related to the variance of payoffs in the Bad Times treatment. Hypothesis 3 posits that the riskiness of payoffs should affect the observed differences between the Good and Bad Times treatments and will be tested *between* subjects.

The chosen parameters values should give sufficiently distinguishable theoretical predictions. To give an idea the model's quantitative predictions, we make additional assumptions. First, since each round typically lasts less than one minute, we assume that, in practice, there is no discounting (i.e., $r = 0$). Second, we assume that the utility function has a power form as in (14). Third, we use the coefficient estimates of Tversky and Kahneman (1992). They estimate that the median exponent of the utility function is 0.88 for both gains and losses (i.e. $\alpha = \beta = 0.88$). The loss aversion coefficient λ is estimated to be equal to 2.25 but, as explained in Section 3.2, it should have no influence on the asset sale decision. For these parameter values, the theoretical predictions for the optimal sale timing are given in Table 2.

Table 2 shows that the distance between the predicted trigger is very substantial in the High payoff parametrization and it decreases with the payoff variance. Within each treatment, the predicted differences between payoff scenarios are less pronounced but still appreciable, especially for the comparison between High and Low.

The experiment was conducted at CentERLab at Tilburg University and our subjects were 100 Tilburg University students recruited using an on-line recruitment process. Participation was voluntary and no individual participated in more than one treatment. Fifty subjects participated in the Good Times treatment and fifty subjects participated in the Bad Times treatment. The nationality and gender composition of the groups were similar. For each treatment, we ran six separate sessions. The experiment was programmed and conducted using z-Tree (Fischbacher, 2007).

At the beginning of each treatment, instructions were read aloud. Each subject played 15 practice rounds. Earnings (initial wealth plus payoff) were paid at the end of the experimental sessions for one of the 45 rounds and the payment round was chosen at random at the end of the experiment. The average earning was 15.74 euros. Sessions typically lasted less than one hour, including reading of instructions and payment.

During the experiment, subjects were seated at isolated computer terminals. At the beginning of each round, the computer screen shows the parameter values. There is a button labeled OK at the bottom-right of the initial screen. By clicking OK, a new screen appears that repeats the same information as the previous screen. Additionally, at the bottom right of the new screen there two buttons: PLAY and STOP. By clicking on STOP, subjects immediately sell the risky assets and obtain the selling price M. By clicking PLAY, subjects begin the sale game and a new screen appears. This screen shows a running timer. At the bottom right of the new screen, there is a button STOP. When subjects want to liquidate the risky asset, they click STOP and the round finishes.

	Number of observations (censored, %)		
	High	Medium	Low
Good Times	729 (239, 33%)	780 (231, 30%)	741 (337, 45%)
Bad Times	779 (465, 60%)	760 (367, 48%)	711 (351, 49%)

Table 3: Number of observations, along with number and percentage of censored observations for the Good and Bad Times treatments in the different payoff schemes.

5 Experimental Results

5.1 Testing the Hypotheses

There are a total of $100 \times 45 = 4,500$ observations, $2,250$ for each treatment. Data are right-censored. When the risky asset is revealed to be of high quality or a negative shock occurs, the exit decision is not observed as the termination is not chosen by the subjects. In the Good Times treatment, subjects failed to sell before expiration for exogenous reasons in 807 (36%) of the cases. In the Bad Times treatment, subjects failed to sell in $1,183$ (52%) of the cases. The number of total observations and of censored observations for each payoff scheme is shown in Table 3.

Table 3 reveals that the proportion of censored observations is larger in the Bad Times treatment. Furthermore, while in the Good Times treatment this proportion is larger for the payoff scenario Low, in the Bad Times treatment it is larger for the payoff scheme High. In general, a larger number of censored observations signals a risk-seeking attitude. Typically, if the subjects have a stronger propensity to take risks, they wait longer and are more likely to fail to sell before an external events occurs. In this respect, the preliminary evidence reported in Table 3 is consistent with the main prediction of the model.

Censored observations cannot be simply discarded as they contain important information. They reveal that subjects were willing to wait at least up until the time when an external event occurred. For this reason, removing censored observations would lead to an underestimation of the exit time. To deal with the problem of censored data, we estimate the subjects' propensity to wait by using the non-parametric product limit estimator (Kaplan

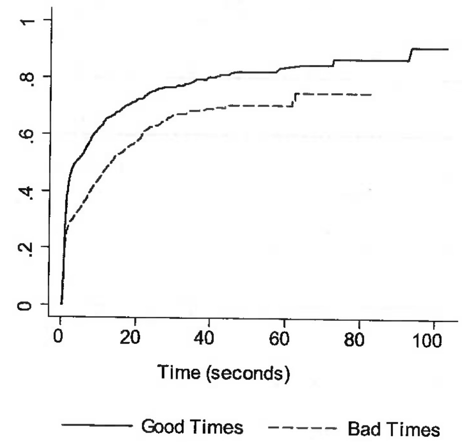


Figure 1: Product limit estimate of CDFs for the sale trigger for the Good and Bad Times treatments.

and Meier, 1958). This method accounts for random censoring and is a standard statistical tool in medical studies in which a number of subjects are lost for exogenous reasons.

The first step of our analysis is to compare the sale decision in Good and Bad Times by using observations pooled across subjects and payoff schemes. That is, we consider observations as i.i.d. and only distinguish between the two treatments. This preliminary investigation provides a first general picture of differences in sale strategies in the gain and loss domains. Figure 1 plots the product limit estimates for cumulative density functions (CDFs) of the sale trigger in Good and Bad Times. For each time (seconds on the x-axis), the figure shows the proportion of cases in which the risky asset was sold. A shift of the curve to the right means that, at each time, the risky asset was sold in a lower number of cases and, therefore, it means that subjects waited longer to sell. The figure shows that the estimated CDF in Bad Times is shifted to the right compared to Good Times. We investigate whether the difference between CDFs is statistically significant by means of a log-rank test. The test rejects the null hypothesis of equality between CDFs ($p = 0.000$). Hence, the general picture that emerges from Figure 1 is consistent with the main prediction of the model and provides support for Hypothesis 1. Subjects sell later in the loss domain than in the gain domain.

We next compare the subjects' sale decisions in Good and Bad Times by dividing data for

	Total	High	Medium	Low
Null rejected (p-value)	Yes*** (0.000)	Yes*** (0.000)	Yes*** (0.000)	No (0.297)

Table 4: Log-rank tests for the between-subjects analysis in the Good and Bad Times treatments. The null hypothesis is the equality between means for the Good and Bad Times treatments in the different payoff schemes. *, **, and *** indicate statistically significant differences at the 10, 5, and 1 percent levels.

the three payoff schemes with varying levels of risk. Figure 2 plots the estimated cumulative density functions (CDFs) of the sale trigger for the Good and Bad Times treatments and for the High, Medium, and Low schemes. The figure shows that for High and Medium, the estimated CDFs in Bad Times is shifted to the right compared to Good Times while CDFs are almost overlapping in Low. The visual impression is confirmed by pairwise log-rank tests. For higher levels of risk (High and Medium), the difference between CDFs is statistically significant ($p = 0.000$ in both cases), while it is not significant in the scheme with the lowest risk (Low with $p = 0.297$). The results again support our main experimental hypothesis on the existence of the disposition-type effect (Hypothesis 1). The results are also consistent with the hypothesized effect of risk on the disposition-type effect (Hypothesis 3). A summary of the between-subjects analysis is found in Table 4.

If the main cause for the observed disposition-type effect is a difference in risk attitude, the evidence up to this point is consistent with differences in relative risk aversion (or relative risk seeking) between the gain and loss domains. Prospect theory says more: individuals are risk averse in the gain domain but risk seeking in the loss domain. If this is the case, also Hypothesis 2 should be supported. To this end, we perform a within-subjects analysis and investigate, for each treatment, differences in sale timing in the three payoff schemes.

Figure 3 presents the CDFs of sale decisions across the payoff schemes. The figure shows that, in Good Times, the CDFs for the parameterizations High and Medium are similar. A pairwise log-rank test cannot reject the null hypothesis of equality of distributions. In contrast, the CDF in the Low parameterization is substantially shifted to the right, which indicates a stronger willingness to wait. A series pairwise log-rank test rejects the null

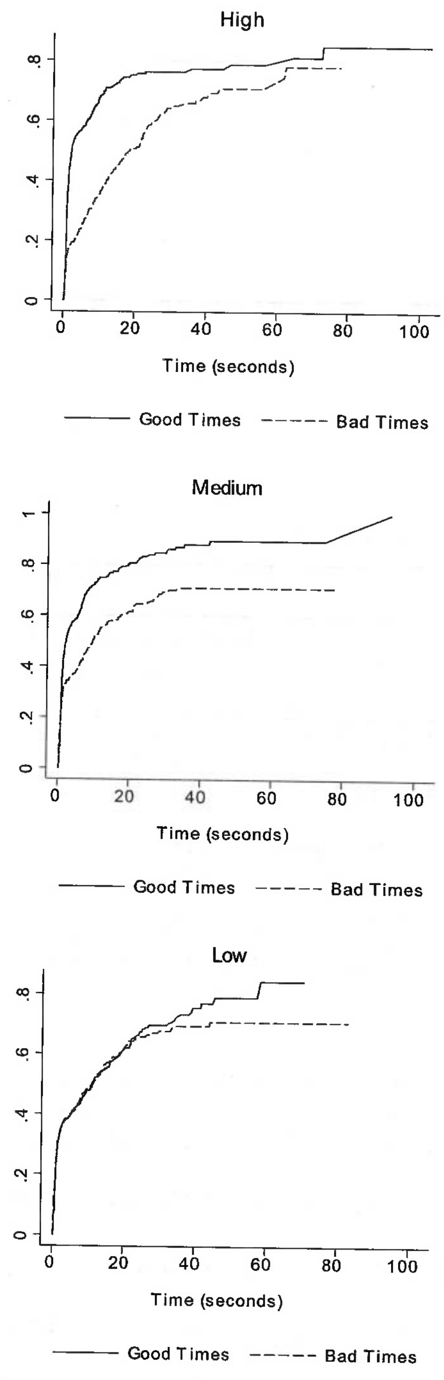


Figure 2: Product-limit estimates of the CDFs of the sale time for the Good and Bad Times treatments in the parameterizations High, Medium and Low.

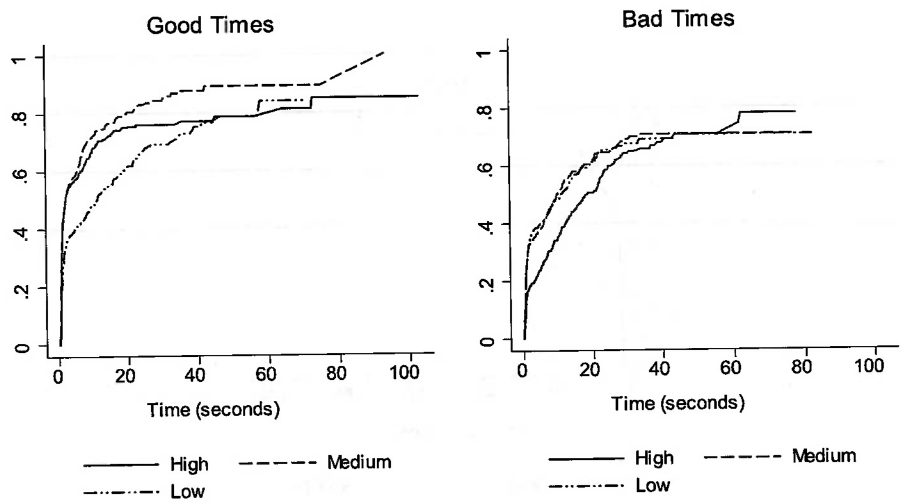


Figure 3: Product-limit estimates of the CDFs of the sale time in the parametrizations High, Medium, and Low for the Good and Bad Times treatments.

	Good Times	Bad Times
	Null rejected (p-value)	
High versus Medium	No (0.252)	Yes*** (0.000)
High versus Low	Yes*** (0.000)	Yes*** (0.000)
Medium versus Low	Yes*** (0.000)	No (0.666)

Table 5: Log-rank tests for the within subjects analysis by payoff scenarios. The null hypothesis is the equality between CDFs. *, **, and *** indicate statistically significant differences at the 10, 5, and 1 percent levels.

hypothesis of equality of the CDFs between Medium and Low, and High and Low. The null is not rejected for the comparison between High and Medium. In contrast, the figure reveals a reverse pattern in Bad Times. While the CDFs are almost overlapping for the Low and Medium parameterizations, the curve for the High parameterization is clearly shifted to the right and its difference with Medium and Low is statistically significant. A summary of the within-subjects analysis is found in Table 5. The pattern is consistent with risk aversion in the gain domain and risk seeking in the loss domain.

The analysis so far relies on the simplifying assumption that observations are i.i.d. How-

	Mean (± Std. Error)			
	Total	**High**	**Medium**	**Low**
Good Times	12.49 (±1.04)	12.24 (±2.03)	8.89 (±1.36)	16.35 (±1.82)
Bad Times	21.26 (±1.64)	25.60 (±2.71)	19.37 (±2.80)	18.82 (±2.98)

Table 6: Product limit estimates and standard errors of the mean of the sale time for the Good and Bad Times treatments in the different payoff schemes.

	Total	**High**	**Medium**	**Low**
Null rejected (p-value)	Yes** (0.037)	Yes*** (0.000)	Yes*** (0.000)	No (0.596)

Table 7: Mann–Whitney test comparing sample means for the by-subject estimates of the sale trigger in the Good and Bad Times treatments. The null hypothesis is the equality between means for the Good and Bad Times treatments in the different payoff scenarios. *, **, and *** indicate statistically significant differences at the 10, 5, and 1 percent levels.

ever, as each subject played 45 rounds, observations are expected be correlated across subjects. To account for within-subject dependence, we construct for each individual a product limit estimate for the average of the sale trigger. Figure 4 shows histograms of the by-subject means for the Good and Bad Times treatments divided by the payoff parametrizations. Evidently, subjects sell the risky asset later in Bad Times and the disposition behavior is more pronounced in High. Furthermore, subjects sell later in Low than in High in the Good Times treatment, while the reverse pattern is observed in the Bad Times treatment.

Table 6 presents the means for the by-subject analysis in the different payout parametrizations. We formally investigate the differences in sale timing with a series of pairwise Mann-Whitney tests. The between-subject analysis for the differences between the gain and loss domains are shown in Table 7. The table shows that, pooling different payoff schemes, the difference in means between Good and Bad Times is statistically significant. Dividing by the payoff parametrizations, the differences are significant in the High and Medium parametrizations but not in the Low parametrization. This is consistent with Hypothesis 3 that the disposition-type effect should be stronger when the payoff variance is large and weaker for low payoff variance.

Table 8 presents a series of pairwise tests for the within-subjects analysis. The High versus Low comparison shows that the differences in extreme payoff parametrizations are

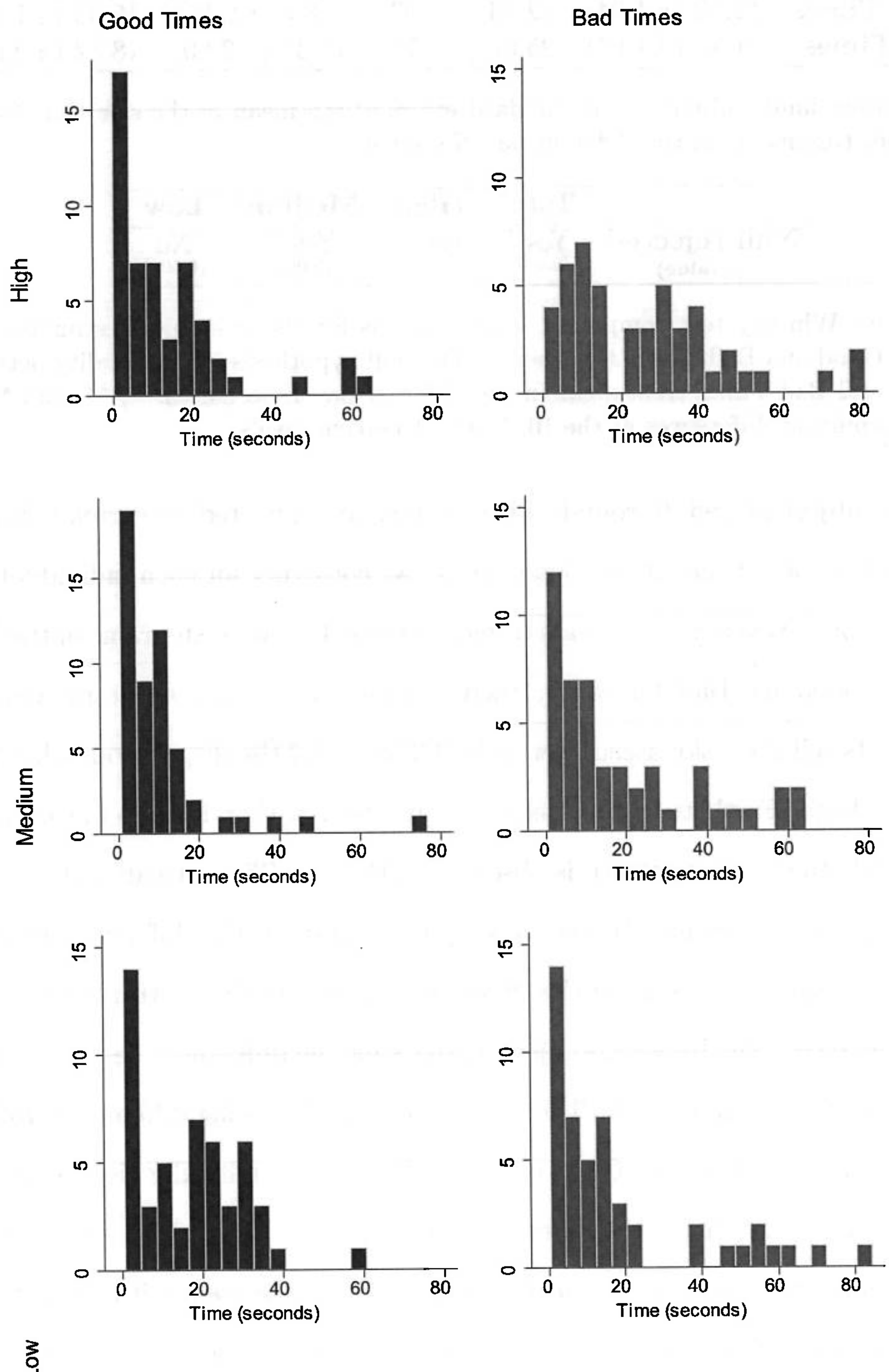


Figure 4: Product limit mean estimates of the sale trigger for the Good and Bad Times treatments in the different payoff schemes.

	Good Times	Bad Times
	Null rejected (p-value)	
High versus Low	Yes** (0.027)	Yes** (0.022)
High versus Medium	No (0.469)	Yes** (0.040)
Medium versus Low	Yes*** (0.000)	No (0.746)

Table 8: Mann–Whitney test comparing the sample means for the by-subject estimates of the sale trigger in the different payoff schemes. *, **, and *** indicate statistically significant differences at the 10, 5, and 1 percent levels.

statistically significant in a way consistent with the reflection effect in preferences and with Hypothesis 2.

5.2 Risk Coefficients Estimation

Exploiting the advantage of our design, we investigated the qualitative predictions of the model without relying on particular functional forms of the utility function. We now turn to a more quantitative analysis of the individuals' timing decisions and estimate implied parameters of risk attitudes. The objective here to determine whether the observed behavior in our experiment is consistent with risk attitudes identified in previous studies. In other words, we try to gauge if the behavior can be reasonably considered consistent with the model. A quantitative mismatch between the subjects' behavior and the model could cast doubt on our qualitative analysis above.

Following Tversky and Kahneman (1992), we assume that individuals' preferences are described by the two-piece power utility function defined in (14). We estimate the following econometric model using nonlinear least squares:

$$t_i^* = \tau_i^* \left(H, M, L, x\right) + \varepsilon_i, \tag{16}$$

where t_i^* is subject i's product limit estimate of the average investment time, $\tau_i^* \left(H, M, L, x\right)$ is the functional form for the optimal timing specified in (11), $x = \{\alpha, \beta\}$ are the risk-

Risk coefficient estimation				
	Coefficient	Std. Error	t-stat	95% Confidence Interval
Good Times	0.48	0.031	15.69	0.42 – 0.54
Bad Times	0.84	0.042	19.95	0.75 – 0.92

Table 9: Estimates of the coefficient of risk aversion in the gain and loss domain.

attitude coefficients of the utility function in the gain and loss domains, and ε_i is an error term assumed to be orthogonal to the average timing decision t_i^*.[10]

Table 9 presents the estimation results. The estimated parameters are $\alpha = 0.48$ and $\beta = 0.84$. The curvature of the utility function in the loss domain is similar to Tversky and Kahneman's (1992) estimate of $\beta = 0.88$. The curvature in the gain domain is more pronounced than Tversky and Kahneman's (1992) estimate of $\alpha = 0.88$. This indicates a higher risk aversion for gains in our data. Our estimate is, however, similar to estimates of α in other elicitations for this utility function; in particular, Wu and Gonzalez (1996) estimate α at 0.50 or 0.48 depending on the method (they do not estimate β). Overall, it seems safe to conclude that the behavior of the experiment subjects is consistent with realistic risk attitudes in Tversky and Kahneman's utility function.

6 Conclusions

In this paper, we examine individuals' timing decisions to sell a risky asset in good and bad times. We present a model in which the asset is subject to large risks. The model predicts that under prospect theory preferences, individuals will wait longer to exit or to sell a risky asset with negative payoffs (bad times) than with positive ones (good times), and that this difference becomes more pronounced when the riskiness of the payoffs increases. Experimental evidence confirms our predictions.

Our design is not expressly about security trading and, therefore, does not settle the dispute on whether or not prospect theory can explain the disposition effect in the stock

[10]Recall that under the specification given in equation (14), the loss aversion parameter λ cancels out and does not affect the liquidation decision.

market. Our liquidation problem, in which an asset is subject to infrequent discontinuous changes in value, can be a more accurate description of different market settings as, for example, real estate or natural resources exploration.

Our model and the design of the experiment are geared, as much as possible, toward a clear identification of the link between prospect theory and the disposition-type effect. In particular, we can exclude the alternative explanation that individuals (mistakenly) believe in the mean reversion of asset prices. The inability to distinguish between the two explanations has often been a confounding factor of previous studies based on both experimental and field data (Weber and Camerer, 1998; Odean, 1998).

Appendix

A Proofs

A.1 Proof of Proposition 1

The result follows directly from applying Assumption 3.1 to equations (7) and (9). ■

A.2 Proof of Proposition 2

Since τ^* in (9) is decreasing in p^*, it suffices to show that $p_G^* \geq p_L^*$. Rewrite P_G and P_L as

$$P_G = \frac{\phi}{\mu}\frac{\frac{r+\phi}{\phi}v(\overline{M}) - v(\overline{L})}{v(\overline{H}) - v(\overline{M})},$$
$$P_L = \frac{\phi}{\mu}\frac{\frac{r+\phi}{\phi}v(\underline{M}) - v(\underline{L})}{v(\underline{H}) - v(\underline{M})}.$$

We have that $\frac{r+\phi}{\phi}v(\overline{M}) - v(\overline{L}) \geq v(\overline{M}) - v(\overline{L})$, because $\frac{r+\phi}{\phi} \geq 1$ and $v(\overline{M}) > 0$. Then using that $v(x)$ is concave for $x \geq 0$, and so for any x and y, $v(y) - v(x) \leq v'(x)(y-x)$, we have that

$$v(\overline{M}) - v(\overline{L}) \geq v'(\overline{M})(\overline{M} - \overline{L}),$$
$$v(\overline{H}) - v(\overline{M}) \leq v'(\overline{M})(\overline{H} - \overline{M}).$$

Thus

$$P_G \geq \frac{\phi}{\mu} \frac{v(\overline{M}) - v(\overline{L})}{v(\overline{H}) - v(\overline{M})} \geq \frac{\phi}{\mu} \frac{\overline{M} - \overline{L}}{\overline{H} - \overline{M}}. \tag{17}$$

Symmetric arguments apply for Bad Times. We have that $\frac{r+\phi}{\phi} v(\underline{M}) - v(\underline{L}) \leq v(\underline{M}) - v(\underline{L})$, because $\frac{r+\phi}{\phi} \geq 1$ and $v(\underline{M}) < 0$. Then using that $v(x)$ is convex for $x < 0$, and so for any x and y, $v(y) - v(x) \geq v'(x)(y - x)$, we have that

$$v(\underline{M}) - v(\underline{L}) \leq v'(\underline{M})(\underline{M} - \underline{L}),$$
$$v(\underline{H}) - v(\underline{M}) \geq v'(\underline{M})(\underline{H} - \underline{M}).$$

Thus

$$P_L \leq \frac{\phi}{\mu} \frac{v(\underline{M}) - v(\underline{L})}{v(\underline{H}) - v(\underline{M})} \leq \frac{\phi}{\mu} \frac{\underline{M} - \underline{L}}{\underline{H} - \underline{M}}. \tag{18}$$

By comparing (18) and (17) under Assumption 3.1, $P_G \geq P_L$ and thus by (10), $p_G^* \geq p_L^*$ as claimed. ■

A.3 Proof of Proposition 3

Part (i) follows from applying the same steps as in the proof of Proposition 2 with $v(x) = x$ and $r > 0$. Part (ii) follows from the same steps as in the proof of Proposition 2 with $v(x)$ strictly concave for $x \geq 0$ and convex for $x < 0$, and $r \geq 0$. In both parts (i) and (ii), $\tau_L^* = \tau_G^*$ only if $\tau_L^* = \tau_G^* = 0$.

To asses the effect of loss aversion, we use the two functional forms. After substituting the power value function (14) into (13), the loss-aversion parameter λ cancels so it plays no role.

After substituting the exponential value function (14) into (12) and (13), the parameters ψ_1 and ψ_2 cancel out. Further, for arbitrary positive γ_1 and γ_2, it holds that $\tau_L^* \geq \tau_G^*$ (this is due to the reflection effect in this functional form and it is covered by (ii)). So the loss aversion restriction, $\psi_1\gamma_1 < \psi_2\gamma_2$, does not play a role for the disposition effect. ■

A.4 Proof of Proposition 4

With $r = 0$, P_G and P_L simplify to

$$P_G = \frac{\phi}{\mu}\frac{v(\overline{M}) - v(\overline{L})}{v(\overline{H}) - v(\overline{M})},$$
$$P_L = \frac{\phi}{\mu}\frac{v(\underline{M}) - v(\underline{L})}{v(\underline{H}) - v(\underline{M})}.$$

To show the claimed effect of risk on timing in the two domains, we show that increased risk increases P_G and decreases P_L.

An increase in risk, as defined in Section 3.2, lowers L and increases H such that M and $\frac{M-L}{H-M}$ stay constant. This implies that an increase of H by ΔH must be paired with a decrease of L by $\frac{M-L}{H-M}\Delta H$. Thus a marginal increase in risk, has the following effect on P_G:

$$\begin{aligned}
&\frac{\phi}{\mu}\frac{\frac{\overline{M}-\overline{L}}{\overline{H}-\overline{M}}v'(\overline{L})(v(\overline{H}) - v(\overline{M})) - \left(v(\overline{M}) - v(\overline{L})\right)v'(\overline{H})}{(v(\overline{H}) - v(\overline{M}))^2} \\
&= \frac{\phi}{\mu}\frac{1}{v(\overline{H}) - v(\overline{M})}\left[\frac{\overline{M}-\overline{L}}{\overline{H}-\overline{M}}v'(\overline{L}) - \frac{v(\overline{M}) - v(\overline{L})}{v(\overline{H}) - v(\overline{M})}v'(\overline{H})\right] \\
&> 0.
\end{aligned}$$

The last inequality follows from the concavity of $v(x)$ if $x > 0$, which implies that $v'(\overline{L}) > v'(\overline{H})$ and $\frac{\overline{M}-\overline{L}}{\overline{H}-\overline{M}} > \frac{v(\overline{M})-v(\overline{L})}{v(\overline{H})-v(\overline{M})}$ and so the term in the square brackets is positive. This proves that increased risk increases P_G and, given the inverse relation between P_G and τ_G^*, accelerates investment in the gain domain.

Symmetric arguments apply to the loss domain. A marginal increase in risk, has the following effect on P_L:

$$
\begin{aligned}
&\frac{\phi}{\mu}\frac{\frac{M-L}{H-M}v'(\underline{L})(v(\underline{H})-v(\underline{M}))-(v(\underline{M})-v(\underline{L}))\,v'(\underline{H})}{(v(\underline{H})-v(\underline{M}))^2}\\
&=\frac{\phi}{\mu}\frac{1}{v(\underline{H})-v(\underline{M})}\left[\frac{M-L}{\underline{H}-\underline{M}}v'(\underline{L})-\frac{v(\underline{M})-v(\underline{L})}{v(\underline{H})-v(\underline{M})}v'(\underline{H})\right]\\
&<0.
\end{aligned}
$$

The term in the square brackets in the second line is negative because $v(x)$ is convex if $x<0$ and so $v'(\underline{L}) < v'(\underline{H})$ and $\frac{M-L}{\underline{H}-\underline{M}} < \frac{v(\underline{M})-v(\underline{L})}{v(\underline{H})-v(\underline{M})}$. Thus increased risk decreases P_L and, given the inverse relation between P_L and τ_L^*, delays investment in the loss domain.

The last statement of the proposition that the disposition-type effect (i.e., the difference between timing in the gain and loss domain under Assumption 3.1) is more pronounced with increased risk, follows as a simple corollary of the first two results. ∎

B Instructions

We report the instructions for the Bad Times treatment. Instructions for the Good Times treatment are analogous with only the payoffs changed.

INSTRUCTIONS

You will participate in an experiment on exit decisions where, depending on your performance, you can win a considerable amount of money.

THE EXPERIMENT

You have an initial endowment equal to 35 Euros and you are the owner of a risky asset. The quality of the asset is unknown. It can be either good or bad. The probability that the risky asset is good is 65%.

If the quality is good, there is a positive probability that the risky asset will give you a high payoff **H**. This probability is equal to **p**=4% per second. If the quality is bad, the probability to obtain the high payoff is zero.

At each instant, a negative shock can occur. If a negative shock occurs you obtain a low payoff **L**. The probability of a negative shock is **q**=4% per second.

To avoid the negative shock, you can stop the game. If you stop the game, you get a safe payoff **M**. The high payoff is larger than the safe payoff which is larger than the low payoff (**H**>**M**>**L**).

WHAT DO YOU HAVE TO DO?

It is a dynamic game. Time goes on and you have to decide whether to stop and get **M** or to wait and try to obtain **H**.

At each instant one of these four scenarios can occur:

1. You receive **H**.

2. You stop to get **M**.

3. A negative shock occurs and you receive **L**.

4. You wait, nothing happens, and the game goes on.

The game stops when either 1, or 2, or 3 occurs. You will play the game for 45 rounds plus 15 practice rounds.

PAYOFFS

In each round, one of the following scenarios can occur (payoffs are in Euros):

Scenario 1

H= 0; **M**= -20; **L**= -30.

Scenario 2

H=-6; **M**= -20; **L**= -27.

Scenario 3

H= -14; **M**= -20; **L**= -23.

Each scenario occurs with probability 1/3 (33.3%). The payoff scenario can change at every round.

SCREEN INFORMATION

At the beginning of each round the initial screen shows the information about the value of the parameters. You find a button "OK" at bottom-right of the initial screen. When you are ready to start, click "OK". If you click "OK" a new screen appears. In the new screen you see the same information as in the previous screen. At bottom-right of the screen you find two buttons: "PLAY" and "STOP". If you want to immediately sell the risky asset and get **M**, click on "STOP". If you want to play, click on "PLAY". If you click "PLAY" a new screen appears. The screen shows a box with the time (seconds) running. At the bottom-right of the screen there is a button "STOP". When you want to sell the risky asset, click on "STOP".

PAYMENT

At the end of the experiment, you will be paid the amount of money (initial wealth plus payoff) that you get in one of the rounds. The payment round is chosen at random at the end of the experiment.

References

Anderson, S.T., D. Friedman, and R. Oprea, 2010. Preemption games: Theory and evidence, *American Economic Review*, 100, 1778–1803.

Barber, B.M., and T. Odean, 2011. The Behavior of Individual Investors. In G. Constantinides, M. Harris, and R. Stulz (eds.), *Handbook of the Economics of Finance.* Amsterdam: Elsevier.

Barberis, N., and W. Xiong, 2009. What drives the disposition effect? An analysis of a long-standing preference-based explanation, *Journal of Finance*, 64, 751–784.

Barberis, N., and W. Xiong, 2012. Realization utility, *Journal of Financial Economics*, 104, 251–71.

Ben-David, I., and D. Hirshleifer, 2012. Are investors really reluctant to realize their losses? Trading responses to past returns and the disposition effect, *Review of Financial Studies*, 25, 2485–2532.

Décamps, J.P., and T. Mariotti, 2004. Investment timing and learning externalities, *Journal of Economic Theory*, 118, 80–102.

Della Seta, M., S. Gryglewicz, and P.M. Kort, 2014. Willingness to wait under risk and ambiguity, Working Paper.

Dixit, A., 1989. Entry and exit decisions under uncertainty, *Journal of Political Economy*, 97, 620–638.

Dumas, B., 1991. Super contact and related optimality conditions, *Journal of Economic Dynamics and Control*, 15, 675–685.

Frazzini, A., 2006. The disposition effect and underreaction to news, *Journal of Finance*, 61, 2017–2046.

Fischbacher U., 2007. z-Tree: Zurich toolbox for ready-made economic experiments, *Experimental Economics*, 10, 171–178.

Frydman, C., N. Barberis, C. Camerer, P. Bossaerts, and A. Rangel, 2014. Using neural data to test a theory of investor behavior: An application to realization utility, *Journal of Finance*, 69, 907-946.

Grinblatt, M., and M. Keloharju, 2001. What makes investors trade? Journal of Finance, 56, 589–616.

Heath, Ch., S. Huddart, and M. Lang, 1999. Psychological factors and stock option exercise, *Quarterly Journal of Economics*, 114, 601–627.

Henderson, V., 2012. Prospect theory, liquidation, and the disposition effect, *Management Science*, 58, 445–460.

Ingersoll, J., and L. Jin, 2013. Realization utility with reference-dependent preferences, *Review of Financial Studies*, 26, 723–767.

Kahneman, D., and A. Tversky, 1979. Prospect theory: An analysis of decision under risk, *Econometrica*, 47, 263–292.

Kaplan, E., and P. Meier, 1958. Nonparametric Estimation from Incomplete Observations, *Journal of the American Statistical Association*, 53, 457–481.

Kaustia, M., 2010. Prospect theory and the disposition effect, *Journal of Financial and Quantitative Analysis*, 45, 1–36.

Köbberling, V., and P.P. Wakker, 2005. An index of loss aversion, *Journal of Economic Theory*, 122, 119–131.

Kyle, A., H.O. Yang, and W. Xiong, 2006. Prospect theory and liquidation decisions, *Journal of Economic Theory*, 129, 273–285.

Li, Y., and L. Yang, 2011. Prospect theory, the disposition effect, and asset prices, *Journal of Financial Economics*, 107, 715–739.

Manso, G., 2011. Motivating innovation, *Journal of Finance*, 66, 1823–1860.

Miao, J., and N. Wang, 2007. Experimentation under uninsurable idiosyncratic risk: An application to entrepreneurial survival, *Columbia University Working Paper.*

Myagkov, M., and C. Plott, 1997. Exchange economies and loss exposure: Experiments exploring prospect theory and competitive equilibria in market environments, *American Economic Review*, 87, 801–28.

Neilson, W.S., 2002. Comparative risk sensitivity with reference-dependent preferences, *Journal of Risk and Uncertainty,* 24, 131–142.

Odean, T., 1998. Are investors reluctant to realize their losses?, *Journal of Finance*, 53, 1775–1798.

Oprea, R., D. Friedman, and S.T. Anderson, 2009. Learning to wait: A laboratory experiment, *Review of Economic Studies*, 76, 1103–1124.

Pastor, L., L. Taylor, and P. Veronesi, 2009. Entrepreneurial learning, the IPO decision, and the post-IPO drop in firm profitability, *Review of Financial Studies*, 22, 3005–3046.

Schwienbacher, A.,2008. Innovation and venture capital exits, *Economic Journal,* 118, 1888–1916.

Shefrin, H., and M. Statman, 1985. The disposition to sell winners too early and ride losers too long: Theory and evidence, *Journal of Finance*, 40, 777–790.

Tversky, A., and D. Kahneman, 1979. Advances in prospect theory: Cumulative representation of uncertainty, *Journal of Risk and Uncertainty*, 5, 297–323.

Wakker, P.P., and A. Tversky, 1993. An axiomatization of cumulative prospect theory, Journal of Risk and Uncertainty, 7, 147–176.

Weber, M., and C. Camerer, 1998. The disposition effect in securities trading: an experimental analysis, *Journal of Economic Behavior and Organization* 33, 167-84.

Weitzman, M.L., 1979. Optimal search for the best alternative, *Econometrica*, 47 641–654.

Wu, G., and R. Gonzalez, 1996. Curvature of the probability weighting function. *Management Science*, 42, 1676–90.

32

Disclosure and agency conflict: Evidence from mutual fund commission bundling

Article *in* Journal of Financial Economics · February 2011

Impact Factor: 3.72 · DOI: 10.1016/j.jfineco.2011.09.007

3 authors, including:

Richard Evans
University of Virginia

Disclosure and Agency Conflict in Delegated Investment Management: Evidence from Mutual Fund Commission Bundling

Roger M. Edelen
Graduate School of Management
University of California, Davis
Davis, CA 95616
rmedelen@ucdavis.edu

Richard Evans
Darden School of Business
University of Virginia
Charlottesville, VA 22906
evansr@darden.virginia.edu

Gregory B. Kadlec
Pamplin College of Business
Virginia Tech
Blacksburg, VA 24060
kadlec@vt.edu

February 1, 2011

Abstract

This study provides empirical evidence on the role of disclosure in resolving agency conflicts in delegated investment management. For certain expenditures fund managers have alternative means of payment which differ greatly in their opacity: payments can be expensed (relatively transparent); or bundled with brokerage commissions (relatively opaque). We find that the return impact of opaque payments is significantly more negative than that of transparent payments. Moreover, we find a differential flow reaction that confirms the opacity of commission bundling. Collectively, our results demonstrate the importance of transparency in addressing agency costs of delegated investment management.

Keywords: Agency Conflict, Mutual Fund; Performance; Brokerage Commissions; Expenses.

*We are grateful for the comments and suggestions of George Aragon, Brad Barber, Michael Goldstein, Paul Irvine, Marc Lipson, Jay Ritter, Michael Schill, Laura Starks, Sunil Wahal, an anonymous referee and seminar participants at the 2009 American Finance Association Meetings, National Bureau of Economic Research Behavioral Finance Conference, University of Oregon Institutional Investor Conference, Arizona State University, University of Arizona, University of Florida, University of Wisconsin at Madison, Virginia Tech and the University of Virginia.

1. Introduction

Agency conflicts have long been a concern with delegated investment management. Indeed, the main impetus behind the 1940 Investment Company Act was the significant potential for agency conflict inherent in delegating management of an investment portfolio, and the role of disclosure in mitigating such conflicts.[1] The literature examines various dimensions of agency conflict at mutual funds, including risk-shifting, market timing, and cross subsidization (see e.g., Brown, Harlow, and Starks, 1996; Zitzewitz, 2003; Gaspar, Massa, and Matos, 2006). Separately, several studies examine how fee disclosure affects investor choice (see e.g., Barber, Odean, and Zheng, 2005; Bergstresser, Chalmers, and Tufano, 2009; Carlin 2009; Anagol and Kim, 2010). However, little is known about the role of disclosure in mitigating agency conflicts in delegated investment management.

This study provides new empirical evidence on the matter by exploiting the fact that fund managers have alternative channels for paying operating costs which differ greatly in opacity. In particular, for certain goods and services managers can bundle their payments with the commissions paid to brokers for trade execution. Bundling payments reduces fund assets with no income statement recognition, in contrast to the alternative of expensing the payment (i.e. including it in the fund's income statement). This reduces transparency on two fronts. First, expensed payments are widely reported via the expense ratio whereas commission payments are reported only in relatively obscure SEC filings. Second, expensed payments are itemized whereas commission payments are not. Therefore a comparison of the performance impact of expensed payments to that of bundled payments for comparable goods and services provides direct insight into the role that disclosure plays in mitigating agency costs.

[1] From the SEC: "The focus of this Act [Investment Company Act of 1940] is on disclosure to the investing public of information about the fund and its investment objectives, as well as on investment company structure and operations." [http://www.sec.gov/about/laws.shtml]

Our central hypothesis is that greater transparency in fund operating expenditures results in lower agency costs, i.e., better return performance. A key requirement for our empirical test is comparability of services received for expensed versus bundled payments. There are three general categories of fund operating costs: advisory (fund manager salary, research, data feeds, transaction cost analytics), administrative (shareholder services, legal, accounting, custodial, transfer agent), and distribution (sale of fund shares). Advisory and administrative costs are difficult to compare across funds because the categories are relatively broad; moreover there can be considerable overlap between the two.[2] By contrast distribution costs are relatively narrow in scope and are governed by statute (rule 12b-1 of the 1940 Investment Company Act).[3] Thus, fund distribution expenditures are uniquely suited to our tests. In particular, comparability ensures that a differential return impact associated with bundled versus expensed payments for distribution can be reasonably attributed to transparency.

To test our hypothesis we obtain expense and brokerage commission data for an exhaustive sample of mutual funds from N-SAR filings with the Securities and Exchange Commission (SEC). Fund disclosure does not itemize commission payments. However, in a large sample commissions can be statistically decomposed into components reflecting payment for trade execution and (bundled) payments for other services by regressing the total commission payment on fund characteristics that affect trade-execution costs. The residual or 'excess' commission from this regression model is our proxy for bundled payments for other services. We combine this proxy for bundled commission payments with additional information from the N-SAR filing to isolate funds' payment for distribution support (i.e., compensation to brokers for selling the

[2] For example, consider bundled payments for research. While the investment-adviser fee might seem an appropriate expense analog, adviser fees include many non-research related costs such as rent, bookkeeping, printing costs, accounting, fund pricing and even trading costs. See Investment Company Institute, 2006, "Mutual Funds and Institutional Accounts: A Comparison", pg. 10; http://www.ici.org/pdf/ppr_06_mf_inst_comparison.pdf .
[3] See Appendix A for a brief regulatory history of commission bundling.

fund), and then analyze the differential return impact of bundled payments for distribution versus expensed payments for distribution. We find strong evidence that transparency helps to mitigate agency costs. In particular, the return impact of the payment is significantly more negative when disclosure is opaque (bundled with commissions) than when it is transparent (expensed). This result is confirmed when we compare all bundled payments to all expenses, and when we compare all non-distribution bundled payments to non-distribution expenses.

One potential criticism is that the use of commission bundling, and more generally the magnitude of funds' commission payments, is plausibly related to other fund characteristics besides trade-execution costs (which we control for) that could generate a spurious relation to returns. For this reason, our analysis includes a wide variety of additional controls relating to other fund, broker, and investor characteristics. We also use an exogenous event – an SEC ruling change in December 2004 that prohibited the use of commission bundling to pay for distribution expenditures – to further establish that our results are not driven by spurious influences. More specifically, we analyze a sample of funds whose commission payments include distribution bundling during part of the sample (pre ban), and don't include distribution bundling during the remainder of the sample (post ban). This implicitly controls for all persistent fund characteristics. Again we find that distribution-related commission payments (occurring pre-ban) are negatively related to performance, reliably more so than distribution expenses. But when those same funds make non-distribution related commission payments (i.e., post ban), those payments are unrelated to returns. This strongly suggests that the return effects we identify are due to the bundling of distribution payments, not to a spurious, omitted fund characteristic.

The second dimension of our analysis examines the impact of disclosure on investor behavior. We find significant differences in investors' response to bundled versus expensed payments. In particular, we find that investor flows are more positively related to bundled payments for distribution than for expensed payments. This result obtains despite the fact that bundled payments are more detrimental to performance. Thus, the agents of investment management – fund managers – appear to garner more benefit from opaque, bundled payments for distribution than from transparent payments. Overall, our analysis demonstrates that opacity in payment disclosure has a substantial effect on investors' response to and monitoring of agency conflicts, and that the magnitude of those conflicts is highest when monitoring is inhibited by opacity.

We are not the first to conjecture that commission bundling obfuscates information relevant to fund investors, or that it represents a less efficient payment mechanism. Industry observers, regulators, and academics have long raised such concerns (Berkowitz and Logue, 1987; Ambachsheer, 1993; Livingston and O'Neal, 1996; Siggelkow, 1999).[4] However, there has been little (if any) direct empirical evidence relating differential transparency associated with commission bundling to either fund performance or investor flow. Our study provides empirical support for the literature's longstanding concern that commission bundling may be fraught with agency costs that are "overlooked" by investors (i.e., do not affect flows).

In some ways the objective in our study is related to that in three preceding papers that look at the differential between hypothetical fund returns constructed from portfolio holdings and actual fund returns. Grinblatt and Titman (1989) pioneer this approach to measuring the performance drag associated with all operations of the fund, observable and not. Wermers (2000)

[4] The literature looks at a variety of other issues regarding commission bundling. Admati & Pfleiderer (1988), Brennan and Chordia (1993), Johnsen (1994), Biais and Germain (2002), and Livne and Trueman (2002) argue that commission bundling is an optimal means of contracting the sale of information. Conrad, Johnson, and Wahal (2001) find that the excess payment associated with commission bundling is not recovered by better trade execution. Goldstein, Irvine, Kandel, and Wiener (2009) find that the persistence of high per-share commissions reflects the bundling of premium services rather than characteristics of the trade (e.g. trade size and stock price). Serafeim (2008) looks at the prohibition of directed brokerage and its impact on the competitive dynamics of the industry between broker-sold and direct-sold funds.

also uses this return differential to separate the value added by managers from trade costs, fund expenses and other frictions. Kacperczyk, Sialm, and Zheng (2008) emphasize that this return gap relates to unobservable actions by fund managers; that it is persistent; and that it predicts fund performance. One interpretation of the evidence in all of these hypothetical-return studies is that economically relevant information is absent from standard fund disclosure. This conclusion is broad in scope, however, encompassing a wide range of unobservable actions including manager skill, trading costs, the performance of the non-equity portion of the portfolio, IPO allocations, transactions costs, handling of flow, and agency costs. Hence, these studies provide comprehensive evidence on the aggregate impact of unobservable actions, but limited inferences regarding the nature and magnitude of agency conflicts, and the role that transparency plays in resolving those conflicts.[5] By contrast, our study is more focused (and narrow) in scope, but that focus yields higher resolution on the specific matter of how disclosure transparency relates to agency conflicts.

Finally, our analysis contributes to the growing body of evidence on disclosure complexity and investor choice. Carlin (2009) models the strategic competitive implications and concludes that funds have an incentive to decrease the transparency of payments and induce payment complexity. The nature of bundled commission payments suggests that they are less transparent and more complex than expensed payments. Our evidence that investor flows are more positively related to bundled distribution payments despite their more negative impact on performance is consistent with this motive.

The remainder of the paper proceeds as follows. Section 2 describes our sample and data sources and presents summary characteristics of the commission data. Section 3 describes the

[5] For example, Kacperczyk, Sialm, and Zheng (2008) find that the overall impact of unobservable actions is positive, indicating that agency conflicts (which are by definition non-positive) are not the dominant factor captured in their experimental design.

expected commission model. Section 4 contains the performance regression results, Section 5 contains the flow results and Section 6 provides a robustness analysis. Section 7 concludes the study.

2. Data and methodology

Our initial sample consists of all U.S. open-end domestic equity mutual funds in the Morningstar database from January 1996 through June 2009. We merge this sample with a database of semi-annual fund N-SAR filings from the SEC.[6] Below we describe these two databases and the variables taken from each.

2.1. *Morningstar data*

The Morningstar database contains data at the share-class level, including monthly fund returns, total net assets (TNA), expense ratios, distribution fees (12b-1), portfolio turnover, fund investment objective, and equity style categories. We identify our sample of domestic equity funds using Morningstar's investment objective code. More specifically, our sample includes all funds with investment objective equal to aggressive growth, growth, growth & income, equity income or small company (objcodes 1-5). This eliminates global, international, balanced and sector funds along with any fixed income funds (objcodes 6-38). We aggregate all share classes for a given fund and remove observations that are missing return, TNA, expense, or turnover data. Finally, we require each fund to have at least 36 months of returns prior to the date of inclusion in the analysis to estimate factor loadings on a 4-factor model (Carhart, 1997).

2.2. *N-SAR data*

[6]Studies that combine N-SAR with CRSP or Morningstar data include Edelen (1999), Reuter (2006), and Christoffersen, Evans and Musto (2010).

Registered investment companies are required by the Investment Company Act of 1940 to file semi-annual N-SAR reports with the SEC. These filings contain detailed information on a wide variety of fund characteristics. The N-SAR form includes 133 numbered questions.[7] The responses to these questions (question number indicated in parentheses) provide information on trading activities for the period, including the total brokerage commissions paid (question 21), the dollar value of purchases (71.A) and sales (71.B) of portfolio securities, and a list of the top ten brokers as measured by the dollar value of commissions paid to each (20). In addition to commissions paid, funds also report if they received fund sales/distribution support (26.A) from brokers as a result of commissions paid. Although compensation for other goods and services is also disclosed in question 26, we focus on distribution support because it alone has a directly comparable expense-ratio counterpart (12b-1 fees).[8]

N-SAR reports are filed at a 'series' level as opposed to a fund or family level. A series consists of one or more funds within a family, generally grouped because of a common date of inception (e.g., merger or creation of a new 'line' of funds). As a result, series often contain a wide variety of fund types, such as equity, bond, and balanced funds. Many N-SAR data items are reported at the individual fund level, but brokerage commissions (21) are reported only in aggregate for all funds in the series. Because of the difficulty in assessing the brokerage commissions paid by each fund in a multi-fund series, much of our analysis focuses on single-fund series. This limits our sample to 45,514 of 179,798 fund-month observations. However, funds in a single-fund series tend to be large, collectively accounting for roughly 50% of total

[7] A list of the questions and sub-questions can be found at http://www.sec.gov/info/edgar/forms/N-SARdoc.htm. In the description of the variables below we identify the N-SAR question and sub-question (e.g., 72.X is the Xth sub-question under question 72) from which the data is collected in parentheses.

[8] N-SAR question 26 requires funds to identify the "Considerations which affected the participation of brokers or dealers or other entities in commissions or other compensation paid on portfolio transactions". Specifically, they are required to indicate whether their decision to use a broker was due to the sale/distribution of fund shares, receipt of research, receipt of quotations, best execution, receipt of telephone line or wire services, the affiliation of the broker/dealer or a commission rebate program.

assets under management. Section 6 contains a comparison between the funds in single-fund and multi-fund series. It repeats the performance analysis on the full sample (both single and multi-fund series) using a proration algorithm to allocate commissions down to the fund level. There is no qualitative difference in results using the full sample, although the noise introduced by proration reduces the precision of estimates.

The Morningstar and N-SAR databases are merged by hand-matching names in Morningstar and in the N-SAR filings. To verify that the mapping is correct we compare monthly purchase and redemption data from Morningstar with the same values from the N-SAR. In results not tabulated, we calculate the correlation between variables that appear in both databases to ensure the data in the N-SAR is accurate including semi-annual returns, expense ratios, fund and family TNA and the percent common stock. The correlations between the variables in the two databases were 0.87 and greater. We apply filters to the Morningstar / N-SAR merged data, removing any observation where the expense ratio exceeds 10%; total commissions exceed 6%; the commission rate exceeds 1%; the trading volume exceeds 2000%; or any fund in the series is missing investment objective information. These five filters combined remove about 6.9% of the sample, leaving 179,798 fund-month observations.

2.3. Sample fund characteristics

Table 1 provides descriptive statistics for the sample, and subsamples based on fund type, investment objective, and style. The sample statistics for 4-factor alphas, expense ratios, 12b-1 fees, and turnover are comparable to previous studies, but the fund and family TNA are larger. This suggests that the single-fund N-SAR requirement tends to pick up relatively large funds and funds from large families. To address any biases this might cause, we include fund and family size control variables in all of our analyses and we repeat the performance analyses for the full sample in Section 6.

Some of the variables in Table 1 are unique to this study, and thus merit further description. *Trade volume* is the fund's total portfolio purchases and sales scaled by TNA which is distinct from portfolio turnover (the minimum of the fund's purchases and sales scaled by TNA). Note that average trade volume (178%) is roughly twice average turnover (82%). *Broker Herfindahl* is the sum of the squared values of the percent of total commissions paid to each of the fund's the top 10 brokers. The maximum value of 1 for this variable indicates that all commissions were paid to a single broker and the minimum value of 0.1 indicates that commissions were equally split among all ten brokers. The mean and median for broker Herfindahl are both closer to the minimum of 0.1 than to the maximum of 1, suggesting that funds generally use a wide range of brokers. This variable captures potential bargaining power of funds in negotiating commissions by concentrating trading to achieve economies of scale.

Institutional% is the percent of the fund's TNA invested in institutional share classes.[9] The literature (e.g. Del Guercio and Tkac, 2002) suggests that institutional investors are more sophisticated than retail investors, thus *institutional%* controls for endogenous determinants of commission rates relating to investor clienteles. *Broker size* is constructed by first assigning a size to each broker (by aggregating the commission dollars received by each broker across all funds each semi-annual period), and then identifying each fund's primary broker (the broker receiving the most commissions for the semi-annual period). This variable captures potential bargaining power of brokers in negotiating commissions via reputation and market power. It

[9] We identify institutional share classes using the Morningstar share class type variable where available, and otherwise searching for "Inst" in the fund name.

also controls for clientele effects, as larger brokers are more likely to provide advice or other ancillary services.

2.4. Commission payments

Table 1 also provides summary statistics of funds' commission payments. The mean (median) annual total commissions are 23 (14) basis points and the mean (median) commission rate is 12 (11) basis points. For a fund trading stocks with an average share price of $35, a commission rate of 12 (11) basis points is equivalent to paying 4.2 (3.9) cents per share. This is in line with Goldstein, Irvine, Kandel, and Wiener (2009) who find that the vast majority of institutional orders are executed at commission rates of either 5 or 6 cents per share in 1999. Finally, Panels B, C and D of Table 1 report characteristics about the sample composition. Specifically, the table reports the percent of index and broker-sold funds (those with a load, a 12b-1 fee or both), the percentage of the sample belonging to different investment objective classifications and the equity style box.

Commission payments likely vary with respect to many fund characteristics, such as family size, fund size, investment strategy, and brokers employed. To infer the effect of payment mechanism on agency costs, it is important to control for these structural determinants of commissions since each could plausibly introduce a spurious relation to returns. Panel A of Table 2 reports average total commissions and commission rates sorted by various fund characteristics. We sort the sample into quartiles based on the variable identified in the row heading (e.g, 4-factor alpha; Fund TNA, etc.) and then calculate the average total commissions and commission rate. For example, sorting on 4-factor alpha, funds in the 2nd quartile had an average 4-factor alpha of -3.88%, and average total commissions of 0.216% of TNA. Consistent with economies of scale, both total commissions and commission rates are declining in fund and family TNA, and the commission rate is declining in trade volume. There is no clear univariate relation between commissions and 4-factor alpha, broker Herfindahl, institutional%, or broker size. Panel B presents average commissions by various fund categories. Actively managed, broker-sold (i.e. load and/or 12b-1) and small cap funds have both higher total commissions and commission rates.

Table 3 presents summary statistics on the percent of funds that make bundled commission payments for distribution (N-SAR 26.A). 26.2% of the observations disclose receipt of distribution support from their brokers.[10] While index funds and directly sold funds (No-Load/No 12b-1) are less likely to receive distribution support than actively managed or broker sold (Load and/or 12b-1) funds, a sizable percentage do: approximately one in seven index funds (15.0%) and one in eight directly sold funds (12.8%) claim to use commissions to pay for distribution support. Market-capitalization category and investment style are largely unrelated to distribution bundling, but the fund's investment objective is related. In particular, aggressive growth funds have a higher tendency to use commissions to pay for distribution support. Larger funds and funds from larger families are also more likely to bundle distribution payments with commissions. While this conceivably reflects a reverse-causality (engaging a broker distribution channel makes a fund/family large), a more likely explanation is a difference in the availability of broker-distribution channels to large and small fund families. Finally less concentrated broker use, and a larger primary broker are also more likely to use commissions to pay for distribution. As noted above, all of these relations are controlled for in our analysis of commissions and fund returns.

[10] Figure 1 shows that over 40% of funds in the sample paid for distribution via bundled commissions in the early part of the sample.

3. Estimating bundled commissions

Our test of the role of disclosure in mitigating agency costs involves a comparison of the return impact of bundled payments and expensed payments for a comparable service (distribution). This section develops the regression-decomposition approach that we use to estimate bundled commission payments. The performance analysis is presented in Section 4.

3.1 Model structure

The objective of the regression decomposition of commissions is twofold. First, as discussed in Section 2, commissions vary for reasons unrelated to bundled payments for other services; such as differential costs of trade execution.[11] If these other determinants of commissions directly impact fund returns and are not controlled for, this spurious relation could cloud our inferences regarding agency conflicts and disclosure transparency. Second, a funds' propensity to make bundled commission payments may correlate with other determinants of returns; such as the investor clientele of the fund. For example, Bergstresser, Chalmers and Tufano (2009) show that funds employing broker distribution channels underperform on a gross basis (i.e. separate from the negative effect of distribution expenses identified by Malkiel, 1995). One might expect that such funds are also more likely to use commissions to effect bundled payments for distribution. Indeed, it is possible that agency costs associated with bundled commission payments are the source of gross (before expenses) underperformance at broker-channel funds (neither Malkiel nor Bergstresser, Chalmers, and Tufano remove these implicit payments). However, the source of underperformance could be unrelated, in which case a

[11] Evidence suggests that trading costs are detrimental to fund performance (see e.g., Edelen, Evans, and Kadlec, 2007). Trading costs also relate to commission payments, both directly and indirectly. 'Difficult' trades, i.e., those with a large price impact, plausibly warrant "higher touch" brokerage services and thus higher commission payments.

spurious correlation between bundled commission payments and returns again clouds our inferences regarding agency conflicts.

To address both issues, we decompose total commissions into a predicted component designed to capture variation in payments relating to trade execution as well as control for confounding factors, and a residual component which is our proxy for variation in bundled payments for other services. The regressors include fund characteristics that affect the cost of trade execution, and fund characteristics that potentially relate to independent determinants of returns. We note that many of the regressors used in constructing predicted commissions likely correlate with bundled payments for distribution, reducing the power of our tests. However, the overriding consideration is ruling out spurious factors. Moreover, orthogonalizing commission payments against these control variables provides a means to evaluate the severity of spurious influences. In particular, the degree to which 'predicted' commissions negatively relate to returns indicates the severity of spurious influences.

3.2 Brokerage commissions model

Table 4 reports coefficient estimates for the predicted commission model. The unit of observation is a semi-annual period for a given fund (7,597 semi-annual N-SAR filings from 765 funds). The dependent variable is total commissions in columns 1 and 2, and the commission rate in columns 3 and 4. The independent variables include an intercept, indicators for the fund's investment style (Value, Blend, Growth) and market capitalization (Large, Mid, Small Cap.), the natural log of fund and family size, trade volume, an indicator for whether or not the fund is indexed, broker Herfindahl, institutional %, an indicator variable for whether the fund was sold through a broker channel or not (Load/12b-1 Fund), and primary broker size. In columns 2 and

4, a fixed effect for each fund family is included. In our later analyses, column 2 is used to decompose total commissions into the predicted and residual components.

The multivariate analysis of Table 4 largely confirm the univariate analysis of Table 2, and are broadly consistent with Livingston and O'Neal (1996) who examine the determinants of brokerage commissions for a sample of 240 U.S. equity funds from 1989 to 1993. Economies of scale are evident with respect to both fund assets and trading volume: The TNA of both the fund and the family is negatively associated with commission rates and total commissions. Higher trading volume correlates with lower commission rates but, not surprisingly, higher total commissions.

The controls for investor clientele and broker also capture significant variation in commissions. Broker Herfindahl has a negative and statistically significant coefficient, indicating that funds that concentrate their trading with fewer brokers pay a lower commission rate and lower total commissions. Broker Size, on the other hand, is positively related to both total commissions and commission rate, consistent with larger brokers negotiating higher commissions via reputation and market power. With respect to investor clientele, broker-sold funds pay higher total commissions and commission rates while index funds have lower commission rates but the relation with total commissions is marginally statistically significant. Finally, Goldstein, Irvine, Kandel, Wiener (2009) show that institutional commission rates are negatively related to price – consistent with a fixed per share commission. The negative relation with lower priced small cap and value stocks are consistent with their findings.

Columns 2 and 4 of Table 4 include a fixed effect for each fund family, which increases the R^2 substantially but has little effect on the other variables. The one exception is the broker-sold (Load/12b-1 Fund) variable, which loses explanatory power. Since distribution channels are typically a family-wide consideration, this result is not surprising. Overall, the results from the commission model (Column 2) are economically intuitive, suggesting that it is a reasonable way to decompose commissions. Moreover, the substantial explanatory power from the various controls (note the R^2 of 76.9%) suggests that the abnormal commissions from this model are orthogonal to plausible sources of spurious relations to returns.

We now turn to the primary focus of the paper, comparing the return effects of bundled commission payments to expensed payments to investigate the role of disclosure in mutual fund agency conflicts.

4. Disclosure and performance

The analyses in this section involve regressing semi-annual four-factor alphas on total commissions (% of fund TNA) in a variety of specifications. We use total commissions rather than commission rates (% of dollars traded) to ensure that returns, excess commission payments, and expenses are all in the same economic units, namely a percent of fund TNA. We use semi-annual returns to parallel the time period covered by the N-SAR filings. Many variables are available at a monthly frequency thus, the regressions use overlapping observations. We compute Newey-West standard errors with six lags.

4.1. Base specification

The base specification is reported in Table 5, column 1. Consistent with the scale effects documented by Chen, Hong, Huang and Kubik (2004), fund size is negatively related to performance while family size is positively related. Consistent with many studies, the expense ratio is negatively related to performance, though not significantly. As generally found in the literature, turnover is not significantly related to returns (e.g., Edelen, Evans, and Kadlec, 2007).

In column 2 we add total commissions to the regression and find that its coefficient is negative and statistically significant.

Specification 3 of Table 5 includes the predicted and residual components of commissions from column 2 of Table 4. The predicted component is negatively related to returns, though not significantly (t-statistic -1.4). By contrast, residual commissions are strongly negatively related to returns (t-statistic -3.1). This suggests that most of the negative relation between total commissions and returns is driven by bundling of payments rather than trading costs or biases due to spurious factors (both of which are proxied with the predicted component).

The p-value at the bottom of Table 5, column 3 compares the expense ratio's impact on performance to that of residual commissions. This is the most basic comparison of the return impact of opaque versus transparent expenditures. The p-value (0.002) indicates that the return impact from bundled (opaque) payments is statistically more negative than the return impact from expensed (transparent) payments. However, two caveats apply to this result. First, these regressors reflect all expenditures. The mix of services purchased via expensing may differ from that purchased via bundling, in which case the comparison is not apples-to-apples, as different services affect returns differently (e.g., Malkiel, 1995). Second, while expenses are measured without error, errors-in-variable is a concern with our proxy for bundled payment. The econometric treatment of errors-in-variable is discussed in Appendix B. The evidence suggests that the tests are conservative – errors-in-variables would bias the tests against finding results.

4.2. Primary specifications

Specification 4 of Table 5 provides an apples-to-apples comparison of the effects of bundling vs. expensing by separating payments into distribution and non-distribution components. Consistent with Malkiel (1995), distribution-related expenses are significantly negatively related to performance, whereas non-distribution related expenses are positively related to performance. Likewise, the impact of residual commissions on performance for funds that bundle distribution is more negative than for funds that do not (coefficient -4.16 vs. -2.12). More importantly, the performance impact of distribution-related residual commissions is substantially more negative than that of distribution expenses (coefficient of -4.16 vs. -1.38, p-value 0.05). As with the comparison of all bundled payments to all expenses, this focused analysis on distribution payments indicates that the relatively opaque payment mechanism (bundled commissions) more negatively relates to return performance. However, in this test comparability in services purchased is assured by rule 12b-1.

Table 5, column 4 also compares the return effects of all bundled commission payments *not* associated with distribution (coefficient = -2.12) to the return effects of non-distribution expenses (coefficient = 0.97). The p-value for the difference (0.002) indicates that opaque payment for services other than distribution also relate more negatively to returns than expensed payments. As in section 4.1, we cannot be sure that the two payment categories involved in this comparison reflect an equivalent mix of services, so this test should be treated with caution. However, because the most egregious category – distribution – has been accounted for, this test likely involves roughly comparable services.

One important issue in interpreting our results is understanding the large negative performance coefficient on residual commissions. This coefficient of -4.16 indicates that for every 1% of TNA paid out in of residual commissions, fund performance decreases by 4.16%. There are two plausible explanations for a greater than parity (i.e. $1 spent on bundled commissions results in a $1 loss) impact on performance. First, it is possible that the large negative coefficient is due to multicollinearity. To ensure that this isn't the case, we perform several checks. The pair-wise correlation of residual commissions with other variables in the regression are all greater than -0.07 and less than 0.13. The condition indices (less than 2.9) and variance inflation factors (less than 2.0) are also well below critical values. Finally, in examining the sequential build-up of explanatory variables in the regression, starting with residual commissions in a univariate specificiation and ending with the full specification, the coefficient estimates for distribution-related residual commissions in the various specifications range from -3.6 to -4.3. All of this suggests that multicollinearity is not an issue.

Second, it is possible that residual commissions are correlated with some other form of agency cost. In addition to the direct (parity) impact of elevated commission payments, funds may engage in excessive trading for the purpose of generating those payments. To the extent that such trades are superfluous, the entire cost the trade (including the bid-ask spread and price impact) is a deadweight drag on performance. Note that estimates place commissions at less than 25% of the total cost of trading.[12] Thus, ancillary trading costs could easily drive a greater-than-parity relation between performance and residual commissions.

We examine this conjecture in Table 5, column 7 by decomposing residual total commissions (residual commission/TNA) into two terms. The first is the residual commission rate (residual commission/total portfolio purchases and sales) times the average percent trade volume (total portfolio purchases and sales/TNA) for all funds matching the fund's investment objective. The second is the fund's percent trade volume divided by the investment-objective average. The first term captures the impact of an elevated commission rate with no confounding influences from 'excess' trading. The second captures that confounding, indirect influence of 'excess' trading.

[12] For example, for a sample of domestic equity mutual funds, Edelen, Evans and Kadlec (2007) estimate that commissions are 0.13% of TNA and that total per-unit trading costs are 0.76%. Industry trade-cost consultants, such as Wagner and Edwards (1993), likewise place commissions well below 10%.

Looking first at the coefficients for funds that bundle distribution (Broker Dist =Yes), we see that the impact of bundled commissions on performance absent the ancillary effect of excess trading is close to parity (-1.63). The coefficient for excess trading is also significantly negative for these funds, consistent with the large negative impact on performance from residual total commissions (columns 4 – 6) reflecting the ancillary costs of excessive trading to effect commission payments. By contrast, for those funds that do not bundle distribution (Broker Dist =No), neither the residual commission rate nor the excess trade volume are negatively related to performance.

Finally, we also consider the possibility that residual commissions proxy for more general agency costs of delegated investment management. To examine this conjecture we include the return gap from Kacperczyk, Sialm, and Zheng (2008) in our regressions (not tabulated). Consistent with their results, we find return gap to be significantly positively related to future performance and including it decreases the significance of predicted commissions, suggesting that return gap proxies for trade costs, among other things. Including return gap, however, does not alter the coefficient or significance of residual total commissions.

4.3. Additional specifications

By construction, residual commissions are orthogonal to all controls provided in Table 4, column 2 (including fund-family fixed effects). Hence, their relation to returns in Table 5 should not be due to these correlated omitted factors. For completeness, in Table 5, column 5 we repeat the analysis of column 4 adding investor, fund family and broker control variables. Column 6 repeats the analysis using total commissions in place of residual commissions. In both cases, including the additional controls subsumes the explanatory power of other variables in the regression (family size and distribution expenses in particular), but strengthen the inferences regarding bundling versus expensing.

4.4. SEC ban on directed brokerage

On December 13th, 2004 the SEC undertook an action that facilitates a powerful exogenous test of the hypothesis that opaque payments for distribution (via commission bundling) are associated with larger performance impairment than transparent (expensed) payments. The SEC imposed a ban on the use of brokerage commissions to pay for distribution.[13] Consider a fund that used bundled commissions to pay for distribution before the ban, but ceased use after the ban. If the incremental return impact of bundled commissions over expensed payments (as identified in Table 5) is attributable to agency costs associated with opacity, then that differential return should cease after the ban. On the other hand, if the incremental return is due to bias in our measure of residual commissions – e.g., a correlation with omitted factors or a link to fund trading costs – then the ban should have little effect on the coefficient estimate since the bias surely obtains both pre and post ban.

Figure 1 plots the percent of funds that used commissions to pay for distribution from 1996 to 2009. Early in the sample period, around 40% of funds bundled distribution payments with commissions, but starting in 2003, it drops dramatically to under 10% by the end of 2006. Given the compliance date of December, 2004 for the ban, it is surprising that not all funds indicate that they no longer bundle distribution support with commissions in 2005. This may be for a couple of reasons. First, while the compliance date is December, 2004, the SEC's formal rule allows funds to make changes to their regulatory filings "at the time of the next regularly scheduled amendment." This could delay the disclosure of the change in practice. Second, anecdotally,

[13] The final rule (17 CFR Part 270) is published in the September 9, 2004 Federal Register, Volume 69, No. 174, pages 54728-54734.

some fund families indicate that while they no longer *pay* for distribution with bundled commissions, they still execute trades through selling brokers with whom the fund has distribution arrangements (i.e. regular expensed 12b-1 payments and loads). Answering the question affirmatively reduces liability.[14]

In Table 6 we repeat the performance analysis of Table 5, limiting the sample to the two years before and two years after the change (2003 to 2006), and to funds that used commissions to pay for distribution prior to the ban but eliminated the practice after the ban. Due to the ambiguity of the end date for bundling, we estimate the regression using both the N-SAR disclosed date (i.e. using the N-SAR filings to identify when the fund did and did not pay for distribution via bundled commissions) and the SEC compliance date. The latter date presumes that the fund ceased the practice of bundling distribution payments on the compliance date even if the N-SAR filing indicates continuation for some time thereafter.

The coefficients and signs on the fund and family size, distribution, and non-distribution control variables from Table 6 are similar to Table 5. The p-value for difference test #1 indicates that the coefficient on residual commissions is statistically different from the coefficient on distribution expenses when payments for distribution are involved (i.e., when the interactive term Broker Dist = Yes). This confirms the inference from Table 5 that opaque (commission bundled) payments for a specific service (distribution) are more costly than transparent payments. Table 6 also lists a p-value comparing the impact of residual commissions on performance before and after the ban (Difference #2). This p-value confirms that residual commissions are strongly negatively related to performance when distribution payments are involved, but when the fund ceases to bundle distribution payments the adverse performance effect disappears.

[14] The SEC allows funds to use selling brokers to execute their trades as long as they have "policies and procedures designed to ensure that its selection of selling brokers for portfolio securities transactions is not influenced by considerations about the sale of funds shares." Federal Register, Vol. 69, No. 174, pg. 54730.

Overall, the results from this event-study analysis of the SEC ban on directed brokerage confirm the validity of our inferences in the general analysis of Table 5: the relatively large negative return impact of residual commissions *vis à vis* that of expenses is associated with opacity in the payment mechanism, not some omitted factor or bias which surely would not have changed coincident with the exogenous action by the SEC.

5. Disclosure and fund flows

Both our evidence and that in Bergstresser, Chalmers and Tufano (2009) shows that distribution payments are performance-eroding. However, we have shown that the performance erosion is greater for the relatively opaque payment mechanism. We argue that this is because the relatively opaque payment mechanism (bundled commissions) fosters greater agency costs. The analysis in this section uses fund flows to substantiate the premise behind this argument – i.e., that bundling payments via commissions is indeed effective in hiding costs.

Table 7 presents the analysis. In column 1 we regress funds' net flow (as a percentage of TNA) on expenses partitioned into distribution and non-distribution components, as well as a variety of controls including a load indicator variable. The results are largely consistent with the existing literature. Flow is strongly positively related to both concurrent aggregate flows (to the fund's investment objective) and the fund's past performance. Funds belonging to a large family have higher net flows, but the marginal effect of being a larger fund is negative. Turnover is unrelated to flow. Finally, while non-distribution expenses are strongly negatively related to future flows, distribution expenses (12b-1 fees) - which are designed to improve flow - are positively related. Thus, the positive effects of marketing on flow more than outweighs the negative effect of disclosing costs (as seen with non-distribution expenses). Note however that funds sold with a load have lower inflows, so the positive effects of marketing do not outweigh the negative effect of disclosure for this salient cost (see e.g., Barber, Odean, and Zheng, 2005).

Table 7 column 2 adds the regressor lagged total commissions, which has a significantly positive impact on net flow. Column 3 considers total commissions partitioned into a predicted and residual component using the base model from Table 4. The coefficient on predicted commissions is negative but not statistically different from zero, consistent with investors deriving little to no performance signal from base commission payments. By contrast, the coefficient on residual commissions is positive and statistically significant suggesting that distribution support associated with bundled payments to brokers is highly effective in generating inflow. The evidence in column 4, where residual commissions are further partitioned using an indicator for distribution support, supports this conjecture. Funds that receive support have significantly higher net inflow, and the point estimate for the coefficient on residual commissions is economically and statistically larger. This suggests that the positive relation between residual commissions and flow is due to the use of bundled payments for distribution support.

Collectively, the evidence in Table 7 strongly indicates that opacity in disclosure is effective in obfuscating costs. The consistently negative relation between non-distribution expenses – which are relatively transparent in disclosure – and flow suggests that investors are generally averse to operating costs. However, when the costs are specifically designed to generate inflow (i.e., distribution payments), the relation is positive – *to the extent that disclosure is opaque*. Thus, the relation to flow remains negative with load payments, which Barber, Odean, and Zheng argue reflects their high salience. The relation swings positive with expensed distribution payments (which are less salient than loads, e.g., often bundled into a single 'expense ratio'). The relation becomes overwhelmingly positive – four times as large as that seen with expensed payments – when disclosure is highly opaque (i.e., bundled commission payments). This provides strong evidence on the important role that transparency plays in investor's assimilation of operating costs in their allocation decisions.

6. Robustness analysis

The preceding analysis is restricted to funds that file their N-SAR reports as a single-fund series. In this section we compare univariate statistics for funds in single-fund and multiple-fund series, and repeat the analyses of Tables 5 and 6 for the combined sample of funds.

Table 8 repeats the univariate analysis of Table 1 for the combined sample. Panels B, C and D indicate that the sample composition is very similar across the single and multi-fund N-SAR series samples. However, fund and family TNA (Panel A) are substantially different across the two samples; the mean (median) fund size of approximately \$3 billion (\$500 million) in the single-fund sample, is almost double the mean (median) of \$1.5 billion (\$288 million) in the full sample. Similarly, the mean and median of the Family TNA in the single-fund sample are roughly double those of the full sample. In addition to fund and family size, there are minor differences in other variables including distribution (12b-1) fees, trade volume and the percent of the fund's TNA in institutional share classes. To address concerns about whether the single-fund sample is representative, we repeat the analysis for the full sample of funds.

Table 9 presents the regression model of commissions. This table is analogous to Table 4, but the aggregate nature of commissions in the multi-fund series leads to two important differences. First, while the regression in Table 4 uses fund-level data, the regression in Table 9 uses series-level data. Because commissions are aggregated at a series level, we aggregate the fund-level independent variables to a series-level. Fund TNA, for example, is replaced by Series TNA. Similarly, the 0/1 index-fund indicator is replaced by the percent of the series in index funds. Second, while Table 4 includes only domestic-equity investment objectives, Table 9 includes balanced, bond and international variables as well. Again, 0/1 investment-objective indicators in Table 4 are replaced by the percent of the series in each investment objective in Table 9 (percentage is determined by fund TNA in columns 1 and 2 and by trade volume in columns 3 and 4).

The results in Table 9 are broadly similar to Table 4. In some cases the statistical significance differs (i.e. the natural log of family TNA), but the coefficient signs and estimates are largely consistent. Looking at the new investment objective classification variables, the results are economically intuitive. International funds pay higher commissions and bond and balanced funds pay lower commissions than domestic equity funds. It is interesting to note that there is an increase in the R^2 of the commission-rate regressions, and a decrease in the total-commission regressions. This fits the evidence from Goldstein, Irvine, Kandel and Wiener (2009) that per-share brokerage commissions are persistent, as follows: the total-commission regressions use fund TNA and the commission-rate regressions use trade volume to assess the percentage of the series in each investment objective. Persistent per-share brokerage commissions imply that commissions should be proportional to trading volume and not fund TNA.

Table 10 repeats the performance regressions of Table 5 for the full sample. These results again mirror the conclusions found in the main analysis. Consider the p-values of the difference test between the proxies for distribution-bundled commissions and expensed commissions. The performance impact associated with the more opaque payment mechanism (bundling) is significantly more negative (both statistically and economically) than that associated with the transparent mechanism (expensing). Similarly, the p-value from the test of bundling vs. expensing non-distribution expenses (Difference Test #2) also confirms the results of the single-fund sample.

Table 11 repeats performance analysis of Table 10, limiting the sample to the two years before and two years after the SEC's ban on the use of brokerage commissions to pay for distribution (2003 to 2006) and to funds that used commissions to pay for distribution prior to the ban but eliminated the practice after the ban. It is analogous to Table 6, with the sole exception that it uses the full sample and not just the single-fund series sample employed earlier in the paper. The results confirm the inferences from the sample drawn from single-fund series (i.e., Table 6): Residual commissions are strongly negatively related to performance when distribution payments are involved, but when the fund ceases to bundle distribution payments the adverse performance effect disappears. In short, although the decomposition of total commissions into a predicted and residual component is less precise with multiple-fund series, the full-sample evidence strongly confirms the results found in the less noisy sample of single-fund series.

7. Conclusion

Disclosure and transparency are central considerations in models of agency costs, yet little is known about the role of transparency in addressing agency conflicts within delegated investment management. This study uses commission bundling as a means to evaluate that role.

Our evidence indicates that fund expenditures are less efficient when paid via the opaque mechanism of commission bundling versus the more transparent mechanism of expensing. One might ask why a manager would utilize such an inefficient mechanism. Our evidence on fund flows suggests an answer: more opaque disclosure of cost means less negative impact on flows. To the extent that payments for distribution erode performance – and they do – fund managers have an incentive to hide these payments from investors. Our evidence shows that obfuscation works.

Our results also suggest that the negative impact of commission bundling on fund performance extends well beyond the direct cost of elevated commission payments. We provide evidence that one ancillary cost is excessive trading to generate payments via commissions. It may also be the case that residual commissions proxy for other forms of agency cost. As we have shown, transparent cost disclosure improves the effectiveness with which shareholders' monitor managers' actions. Thus, the decision to effect payments via residual commissions versus expensing is likely endogenous to the manager's attention to agency costs – poor stewards likely select opaque payment mechanisms. Consistent with this view, Dimmock and Gerken (2009) show that the use of bundled commissions/soft dollars by registered investment advisors is a strong predictor of future fraud. Additionally, a 1998 SEC report finds that 35% of broker-dealers examined had used bundled commissions to pay for items unrelated to either distribution or research and thus, were not allowed by law. The alternative uses included: rent, office furniture, electric bills, personal travel and dining expenses, limousine service, psychology training, concert tickets and more.[15] Thus, the use of commissions bundling may be a proxy for misuse of investor assets.

Finally, an observation: Regulators' views on the use of commissions as a means to pay for distribution services tend to ebb and flow (see section A.1 of the Appendix). As of 2004, the practice is banned. However, the objective of this study is to empirically investigate the importance of disclosure in delegated investment management. Our interest in commission

[15] In particular, section IV.B of the report: http://www.sec.gov/news/studies/softdolr.htm#bd.

bundling – particularly as it relates to distribution – arises because it provides a unique laboratory to achieve our objective. Thus, the fact that this data is historical rather than ongoing is largely irrelevant to the objective of this study. While bundled payment for distribution is now banned, other forms of agency conflict in delegated investment management surely continue to exist. Our study provides some of the first direct evidence that disclosure is fundamental to addressing those costs.

Appendix A: Regulatory History of Commission Bundling

Prior to the shift from fixed to competitive commission rates in 1975[16], bundling of research services was the norm. In a 1974 address, SEC Chairman Ray Garrett, Jr. characterized these practices by saying:

> "The unfixing of commission rates does not require an unbundling of services. It permits it. As a result, portfolio managers will have a choice whether to pay for research by means of portfolio commission dollars ('soft dollars") or with money no designated as brokerage commissions ("hard dollars")."[17]

In sharp contrast to the bundling of research, bundling of distribution was illegal in the 1970s, prohibited by the NASD Anti-Reciprocal rule.[18] In supporting the NASD's rule, which prohibits "reciprocal sales practices", namely allocating brokerage commissions to a given broker in exchange for sales of the fund's shares and vice-versa, the SEC argued that "The reciprocal use of portfolio brokerage has been viewed by the Commission as creating hidden influences behind recommendations to customers" and "inducing improper portfolio management practices"[19] including excessive trading and commission rates for the purpose of selling fund shares. In 1981, however, the SEC reversed itself, adopting rule 12b-1 which "would permit use of fund assets for distribution"[20] under the assumption that the economies of scale associated with attracting additional fund shareholders would benefit existing shareholders. While rule 12b-1 is commonly recognized as the regulation that allows funds to include in the

---

[16] The SEC adopted Rule 19b-3 on January 23, 1975. This rule eliminated fixed commissions on exchange transactions effective May 1, 1975.

[17] Garrett Jr., Ray, June 11, 1974, "Commission Rates – The World Ahead", address to NYSE conference, text available at: http://c0403731.cdn.cloudfiles.rackspacecloud.com/collection/papers/1970/1974_0611_GarrettNYSET.pdf

[18] An excellent review of all of the legal issues surrounding retail mutual fund broker compensation can be found in Krawczyk (2004).

[19] Annual Report of the Securities and Exchange Commission for the fiscal year ended June 30, 1975, pg. 262.

[20] 45th Annual Report of the Securities and Exchange Commission for the fiscal year ended, September 30, 1979, pg. 71.

expense ratio a separate fee used to pay for fund distribution, it also allowed funds to bundle payments for distribution with their commissions.

Appendix B: Errors-in Variables

To see how errors-in-variables could affect the analysis, suppose that commissions are comprised of three terms (asterisks denote true quantities):

$$C = Base^* + Trad^* + Bund^* + u.$$

where $Base^*$ relates to general characteristics like size; $Trad^*$ reflects higher payments for difficult-to-trade assets like small-cap stocks; and $Bund^*$ reflects payments for other services; and u is an orthogonal disturbance term. Write fund returns as

$$Return = \alpha + \beta_T \cdot Trad^* + \beta_B \cdot Bund^* + \beta_E \cdot Exp^* + \varepsilon$$

where Exp^* is expensed payments. Our central hypothesis is that $\beta_B < \beta_E$. Consider regressing

$$Return = a + b_B \cdot Bund + b_E \cdot Exp^* + e,$$

where $$Bund = C - Base^* = Bund^* + (Trad^* + u) = Bund^* + U$$

is a noisy measure of $Bund^*$ (note that expenses are measured without error). If U is pure noise (uncorrelated with returns), then b_B as an estimate of β_B is biased towards zero.[21] This does not distort inferences from our tests (if $b_B - b_E$ is reliably less than zero then so is $\beta_B - \beta_E$). However, if U is correlated with returns, then the bias in b_B could go either direction. We address this latter errors-in-variable concern with three econometric techniques.

First, we use an expansive set of instruments (i.e., fund and family scale, trading costs, investor clienteles relating to advice, broker type, etc.) to proxy for U. While Table 4 indicates that these instruments capture significant variation in commissions relating to omitted factors (R^2

[21] Page 377 in Greene, W., 2000. Econometric Analysis, 4th Edition. Prentice Hall, New Jersey.

of 76.9%), we find no evidence in Table 5 to suggest that U correlates with returns. In particular, the coefficient on predicted commissions – i.e. the aggregation of all instruments for other factors – is not statistically significant. Second, in columns 5 and 6 of Table 5, we directly add these instruments ("Other Controls = Yes"), as well as fund-family fixed effects, to the returns regression to further account for spurious sources of correlation. These changes have virtually no impact on the coefficient estimate for bundled distribution payments. Third, our analysis of the 2004 ban discussed in section 4.4 controls for any persistent fund characteristics, alleviating concerns about a potential errors-in-variables problem. By focusing on the performance impact of the same fund's payment mechanism before and after the ban, instead of comparing funds that bundle with those that don't, this analysis suggests that the results are due to opacity in the payment mechanism, and not omitted factors which likely would not have changed coincident with the exogenous action by the SEC. Collectively, these results imply that U is not significantly correlated with returns and that errors-in-variables biases b_B towards zero.

References

Admati, A., Pfleiderer, P., 1988. Selling and trading on information in financial markets. American Economic Review 78, 96-103.

Ambachsheer, K., 1993. The soft dollar question: what is the answer? Financial Analyst Journal 49, 8-10.

Anagol, S., Kim, H., 2010. The impact of shrouded fees: evidence from a natural experiment in the Indian mutual funds market. Unpublished working paper. Wharton School of Business, University of Pennsylvania.

Biais, B., Germain, L., 2002. Incentive compatible contracts for the sale of information. Review of Financial Studies 15, 987-1003.

Barber, B., Odean, T., Zheng, L, 2005. Out of sight, out of mind: the effects of expenses on mutual fund flows. Journal of Business 78, 2095-2119.

Bergstresser, D., Chalmers, J., Tufano, P., 2009. Assessing the costs and benefits of brokers in the mutual fund industry. Review of Financial Studies 22, 4129-4156.

Berkowitz, S., Logue, D., 1987. The portfolio turnover explosion explored. Journal of Portfolio Management 13, 38-45.

Brennan, M., Chordia, T., 1993. Brokerage commission schedules. Journal of Finance 48, 1379-1402.

Brown, K., Harlow, W., Starks, L., 1996. Of tournaments and temptations: An analysis of managerial incentives in the mutual fund industry. Journal of Finance 51, 85-110.

Carhart, M., 1997. On persistence in mutual fund performance. Journal of Finance 52, 57-82.

Carlin, B., 2009. Strategic price complexity in retail financial markets. Journal of Financial Economics 91, 278-287.

Christoffersen, S., Evans, R., Musto, D., 2010. What do consumers' fund flows maximize? evidence from their brokers' incentives. Unpublished working paper. Darden School of Business, University of Virginia.

Chen, J., Hong, H., Huang, M., Kubik, J., 2004. Does fund size erode mutual fund performance? The role of liquidity and organization. American Economic Review 94, 1276-1302.

Conrad J., Johnson, K., Wahal, S., 2001. Institutional trading and soft dollars. Journal of Finance 56, 397-416.

Del Guercio, D., Tkac, P., 2002. The determinants of the flow of funds of managed portfolios: mutual funds vs. pension funds. Journal of Financial and Quantitative Analysis 37, 523-557.

Dimmock, S., Gerken, W., 2009. Finding Bernie Madoff: detecting fraud by investment managers. Unpublished working paper. Michigan State University.

Edelen, R., 1999. Investor flows and the assessed performance of open-end mutual funds. Journal of Financial Economics 53, 439-466.

Edelen, R., Evans, R., Kadlec, G., 2007. Scale effects in mutual fund performance: the role of trading costs. Unpublished working paper. Darden School of Business, University of Virginia.

Gaspar, M., Massimo, M., Matos, P., 2006. Favoritism in mutual fund families? Evidence on strategic cross-fund subsidization. Journal of Finance 61, 73-104.

Goldstein, M., Irvine, P., Kandel, E., Weiner, Z., 2009. Brokerage commissions and institutional trading patterns. Review of Financial Studies 22, 5175-5212.

Grinblatt, M., Titman, S., 1989. Mutual fund performance: an analysis of quarterly portfolio holdings. Journal of Business 62, 393-416.

Johnsen, B., 1994. Property rights to investment research: The agency cost to soft dollar brokerage. Yale Journal of Regulation 11, 75-113.

Kacperczyk, M., Sialm, C., Zheng, L., 2008. Unobserved actions of mutual funds. Review of Financial Studies 21, 2379-2416.

Krawczyk, S., 2004. Compensation practices for retail sales of mutual funds. The Journal of Investment Compliance Spring, 27-45.

Livingston, M., O'Neal, E., 1996. Mutual fund brokerage commissions. Journal of Financial Research 19, 273-292.

Livne, G., Trueman, B., 2002. The impact of soft dollars on market equilibrium and investor profits. Unpublished working paper. London School of Business.

Malkiel, B., 1995. Returns from investing in equity mutual funds 1971 to 1991. Journal of Finance 50, 549-572.

Reuter, J., 2006. Are IPO allocations for sale: evidence from mutual funds. Journal of Finance 61, 2289-2324.

Serafeim, G., 2008. Directed brokerage no more: the effects of new regulation in the mutual fund industry. Unpublished working paper. Harvard University.

Siggelkow, N., 1999. Expense shifting: an empirical study of agency costs in the mutual fund industry. Unpublished working paper. Wharton School of Business, University of Pennsylvania.

Wagner, W., Edwards, M., 1993. Best execution. Financial Analysts Journal 49, 65-71.

Wermers, Russ, 2000. Mutual fund performance: an empirical decomposition into stock-picking talent, style,transactions costs, and expenses. Journal of Finance 55, 1655-1695.

Zitzewtiz, E. 2003. Who cares about shareholders? arbitrage-proofing mutual funds. Journal of Law, Economics, and Organization, October 2003, 245-280.

Figure 1 – Excess Commissions and Fund Sales/Distribution Support

Figure 1 shows the percentage of sample funds that bundle fund sales/distribution support with brokerage commissions. The responses are taken from question 26.A in the funds' semi-annual N-SAR filing and the percentages are calculated for every semi-annual period (January to June and July to December).

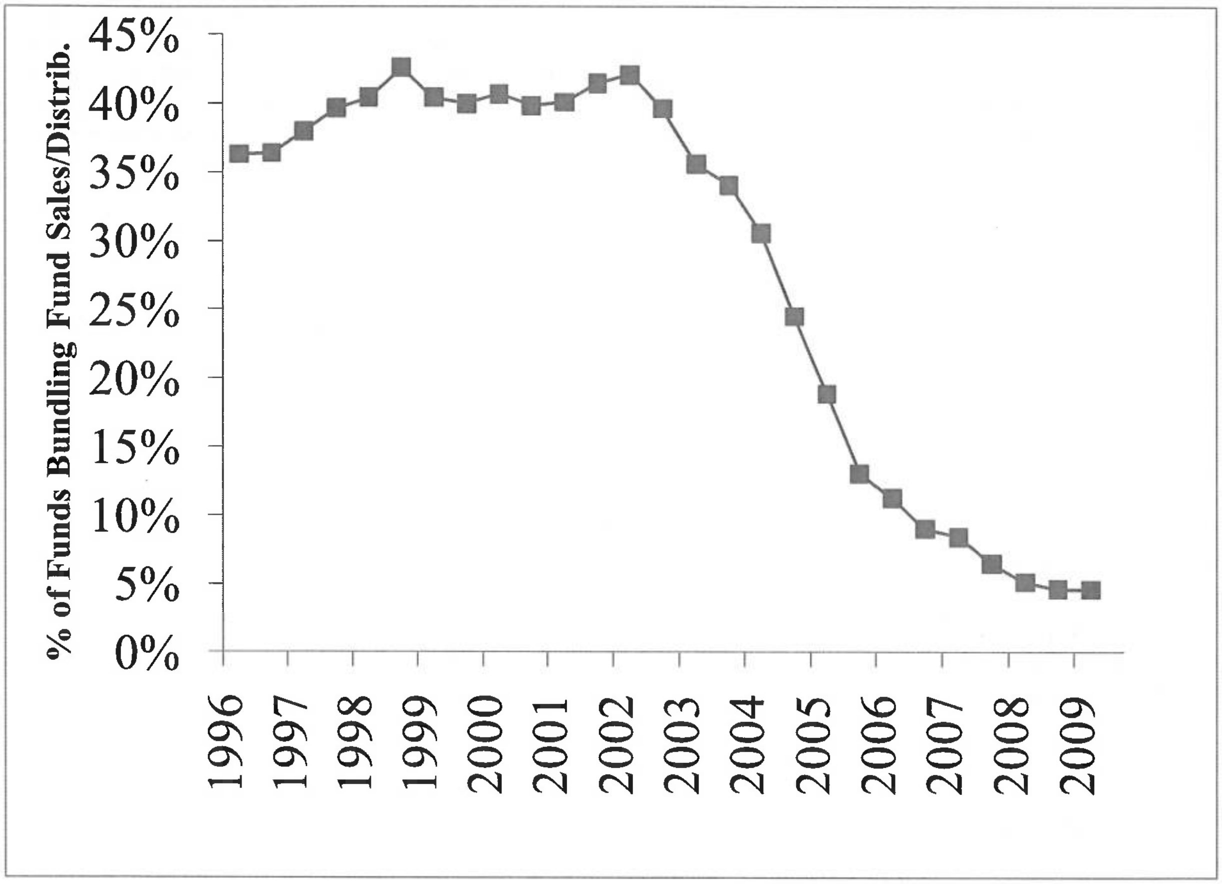


Table 1 – Fund-Level Sample Characteristics

Table 1 presents descriptive statistics for the sample of domestic equity funds with single-fund N-SAR filings over the period January 1996 through June 2009 (45,514 monthly fund observations). The variables include annual 4-factor alpha [Carhart (1997)] using factor loadings estimated over the prior 36 months of data (t-1 to t-36), fund TNA ($millions), family TNA ($billions), annual expense ratio, annual 12b-1 fees for those funds that charge them (27,968 observations), annual turnover (the minimum of fund purchases and sales divided by TNA), annual trade volume (fund purchases plus fund sales divided by TNA), a Herfindahl index of broker payment concentration (calculated using the dollar amount of commissions paid by each fund to their 10 highest compensated brokers, i.e., the maximum value of 1 means that all commissions were paid to a single broker and the minimum value of 0.1 means that commissions were evenly split among the 10 brokers), the percent of the fund's TNA in institutional share classes, broker size (broker size is the log of the annual total commissions received by the fund's primary broker), total commissions (annual brokerage commissions divided by TNA) and commission rates (brokerage commissions divided by the dollar value of purchases and sales over the same period). Panel B reports the number of observations by fund type, specifically the number of observations from index funds and broker-sold funds (i.e. funds that charge a load or a 12b-1 fee or both). Panel C reports the number of observations by Morningstar's investment objective and Panel D reports the number of observations and percent of the sample by the Morningstar style box.

Panel A. Univariate Statistics

Variable	Mean	Median	Percentiles 25th	75th
4-Factor Alpha (Annual)	-1.19%	-1.42%	-6.60%	3.89%
Fund TNA ($MM)	2,905.7	497.9	106.7	1,867.5
Family TNA ($BB)	85.8	19.0	1.2	65.2
Expense Ratio	1.36%	1.23%	0.93%	1.60%
12b-1 Fees (when > 0)	0.41%	0.35%	0.23%	0.54%
Turnover (%)	82%	55%	27%	102%
Trade Volume (% TNA)	178%	115%	58%	216%
Broker Commissions Herfindal	0.24	0.13	0.11	0.24
Institutional % of TNA	9.9%	0.0%	0.0%	3.8%
Broker Size (Annual Commiss in $MM)	138.3	56.4	7.4	300.6
Total Commissions (% TNA)	0.23%	0.14%	0.07%	0.26%
Commission Rate (% Trade Volume)	0.12%	0.11%	0.07%	0.15%

Panel B. Fund Type

	Obs.	Percent
Index Fund	2,177	4.8%
Broker Sold (Load/12b-1)	29,856	65.6%

Panel C. Investment Objectives

	Obs.	Percent
Aggressive Growth	2,590	5.69%
Growth	23,498	51.63%
Growth & Income	9,635	21.17%
Equity Income	2,395	5.26%
Small Company	7,396	16.25%

Panel D. Equity Style Box

	Value	Blend	Growth
Large Cap.	8,075	10,174	8,953
	17.74%	22.35%	19.67%
Mid Cap.	2,149	2,977	5,895
	4.72%	6.54%	12.95%
Small Cap.	1,932	2,251	3,108
	4.24%	4.95%	6.83%

Table 2 – Brokerage Commission Statistics

Table 2 contains summary statistics for total commissions (annual brokerage commissions divided by TNA) and commission rates (brokerage commissions divided by the dollar value of purchases and sales over the same period). Panel A reports statistics for the sample sorted into quartiles based on the variables indicated in the row headings: 4-Factor Alpha [Carhart (1997)] using factor loadings estimated over the prior 36 months of data (t-1 to t-36), fund TNA (in $millions) family TNA (in $billions), annual trade volume (the dollar value of fund purchases and sales divided by TNA). Broker Herfindahl measures how concentrated the fund's trading is across brokers. Institutional % of TNA measures the percent of the fund's TNA in institutional share classes and Log(Broker Size) measures the natural log of the total commissions received from all funds by the fund's primary broker. Panel B reports total commissions and the commission rate when the sample is split by fund type (actively managed vs. index fund or Load and/or 12b-1 vs No Load/No 12b-1), the Morningstar designation of the market cap (Large, Mid, Small) and fund investment style (Value, Blend, Growth) and the fund's investment objective.

Panel A. Independent Variable Quartile Sorts

	Averages by quartile, sorting on variable in bold:			
	1	2	3	4
4-Factor Alpha	-14.72%	-3.88%	1.05%	12.78%
Total Comm. (% TNA)	0.266%	0.216%	0.204%	0.232%
Comm. Rate (% Trade Vol)	0.127%	0.123%	0.120%	0.123%
Fund TNA ($MM)	49.2	301.9	1169.6	10102.3
Total Comm. (% TNA)	0.317%	0.237%	0.213%	0.151%
Comm. Rate (% Trade Vol)	0.144%	0.129%	0.114%	0.105%
Family TNA ($BB)	0.37	7.59	39.31	295.86
Total Comm. (% TNA)	0.306%	0.221%	0.214%	0.179%
Comm. Rate (% Trade Vol)	0.152%	0.123%	0.115%	0.102%
Trade Volume (% TNA)	35%	88%	162%	428%
Total Comm. (% TNA)	0.087%	0.137%	0.224%	0.470%
Comm. Rate (% Trade Vol)	0.131%	0.121%	0.122%	0.118%
Broker Herfindahl	11%	12%	17%	57%
Total Comm. (% TNA)	0.236%	0.217%	0.237%	0.228%
Comm. Rate (% Trade Vol)	0.124%	0.123%	0.126%	0.119%
Institutional % of TNA	0%	0.0005%	2.69%	38.5%
Total Comm. (% TNA)	0.246%	0.240%	0.210%	0.218%
Comm. Rate (% Trade Vol)	0.142%	0.124%	0.120%	0.117%
Log(Broker Size)	6.3	9.3	10.9	12.2
Total Comm. (% TNA)	0.230%	0.241%	0.225%	0.229%
Comm. Rate (% Trade Vol)	0.138%	0.118%	0.122%	0.115%

Panel B. Category Sorts

		Commissions	
		Total (% TNA)	Rate (% Trade Vol)
Fund Type	Indexed	0.030%	0.045%
	Active	0.240%	0.127%
Distribution Channel	Load and/or 12b-1	0.241%	0.125%
	No-Load/No 12b-1	0.212%	0.119%
Market Cap	Large Cap	0.187%	0.112%
	Mid Cap	0.278%	0.131%
	Small Cap	0.316%	0.151%
Fund Investment Style	Value	0.191%	0.133%
	Blend	0.210%	0.117%
	Growth	0.273%	0.122%
Invest. Obj.	Aggressive Growth	0.303%	0.119%
	Growth	0.254%	0.124%
	Growth & Income	0.176%	0.114%
	Equity Income	0.132%	0.108%
	Small Company	0.227%	0.136%

Table 3 – Commissions and Fund Sales/Distribution Bundling

Table 3 reports the number of observations, percent of observations, and percent of funds answering yes to N-SAR question 26.A regarding whether the fund received fund sales/distribution support as a result of their brokerage business. The sample includes all domestic equity funds with single-fund N-SAR filings over the period January 1996 through June 2009 (45,514 monthly fund observations). Statistics are reported for the overall sample, and sub-samples based on fund type (index fund vs. actively managed), distribution channel (Load and/or 12b-1 vs. No-Load/No 12b-1), market capitalization, investment style (book-to-market), investment objective, fund size, family size, trading volume, broker Herfindahl, institutional % of TNA and broker size. The median values for "above median" and "below median" are determined by year and by investment objective.

		Number of Observations	Percent of Sample	Bundled Fund Sales/ Distribution
Full Sample		45,514	100.0%	26.2%
Fund Type	Indexed	2,164	4.8%	15.0%
	Active	43,350	95.2%	26.8%
Distribution Channel	Load and/or 12b-1	29,856	65.6%	33.5%
	No-Load/No 12b-1	15,658	34.4%	12.8%
Market Cap	Large Cap	27,202	59.8%	27.5%
	Mid Cap	11,021	24.2%	25.8%
	Small Cap	7,291	16.0%	22.2%
Fund Investment Style	Value	12,156	26.7%	30.8%
	Blend	15,402	33.8%	24.1%
	Growth	17,956	39.5%	25.0%
Investment Objective	Aggressive Growth	2,590	5.7%	38.6%
	Growth	23,498	51.6%	25.8%
	Growth & Income	9,635	21.2%	27.9%
	Equity Income	2,395	5.3%	27.8%
	Small Company	7,396	16.2%	20.4%
Fund Size	Above Median	22,762	50.0%	33.7%
	Below Median	22,752	50.0%	18.7%
Family Size	Above Median	22,762	50.0%	39.6%
	Below Median	22,752	50.0%	12.9%
Trading Volume	Above Median	22,773	50.0%	31.4%
	Below Median	22,741	50.0%	21.1%
Broker Herfindahl	Above Median	22,750	50.0%	19.7%
	Below Median	22,764	50.0%	33.2%
Inst. % of TNA	Above Median	16,890	37.1%	27.1%
	Below Median	28,624	62.9%	26.0%
Broker Size (Commiss $MM)	Above Median	22,846	50.2%	31.5%
	Below Median	22,668	49.8%	21.3%

Table 4 – Brokerage Commission Model

Table 4 presents coefficient estimates and t-statistics from pooled regressions of brokerage commissions on fund characteristics. The sample includes domestic equity funds with single-fund N-SAR filings over the period January 1996 through June 2009. The observations are semi-annual (7,597 observations) to match the reporting of the N-SAR. The dependent variable is total commissions (brokerage commissions divided by TNA) in specifications 1 and 2 and commission rates (brokerage commissions divided by the dollar value of purchases and sales over the six month period) in specifications 3 and 4. The independent variables include an intercept, the Morningstar designation of fund investment style (Value, Blend, Growth) and market cap (Large, Mid, Small), the natural log of fund TNA and family TNA, trading volume (fund purchases plus sales divided by fund TNA), an index fund indicator, a Herfindahl index of broker payment concentration (calculated using the dollar amount of commissions paid by each fund to their 10 highest compensated brokers, i.e., the maximum value of 1 means that all commissions were paid to a single broker and the minimum value of 0.1 means that commissions were evenly split among the 10 brokers), the percent of the fund in institutional share classes, a load and/or 12b-1 fund indicator, and broker size (the natural log of the total commissions received from all funds by the fund's primary broker). Fund Investment Style and Market Cap Category are indicator variables where Growth and Small Cap are the respective control variables. Specifications (2 and 4) also include fixed effects for fund family. The standard errors are clustered by fund and the resulting t-statistics are provided in parentheses.

	Total Commissions (% TNA)				Commission Rates (% Trade Volume)			
Regression:	1		2		3		4	
Intercept	0.0383	(8.4)			0.3811	(14.8)		
Fund Investment Style								
Value	0.0013	(1.7)	0.0015	(2.3)	0.0144	(3.3)	0.0139	(3.8)
Blend	-0.0003	(-0.4)	0.0008	(1.3)	0.0054	(1.4)	0.0086	(3.0)
Growth	=Control				=Control			
Market Cap. Category								
Large Cap.	-0.0067	(-5.1)	-0.0054	(-5.2)	-0.0479	(-7.6)	-0.0355	(-6.7)
Mid Cap.	-0.0033	(-2.3)	-0.0033	(-3.0)	-0.0266	(-4.1)	-0.0198	(-3.6)
Small Cap.	=Control				=Control			
Log(Fund TNA)	-0.0008	(-3.7)	-0.0010	(-4.7)	-0.0062	(-4.3)	-0.0057	(-4.9)
Log(Family TNA)	-0.0008	(-4.8)	-0.0005	(-3.3)	-0.0046	(-4.3)	-0.0037	(-4.6)
Trading Volume	0.0165	(21.9)	0.0159	(26.9)	-0.0100	(-5.3)	-0.0111	(-7.4)
Index Fund (1=Yes)	-0.0010	(-0.8)	-0.0018	(-1.4)	-0.0601	(-8.6)	-0.0612	(-8.7)
Broker Herfindahl	-0.0057	(-2.1)	-0.0071	(-3.0)	-0.0354	(-2.7)	-0.0296	(-2.5)
Institutional % of Fund	-0.0007	(-0.6)	0.0006	(0.4)	-0.0031	(-0.4)	0.0021	(0.3)
Load/12b1 Fund (=Yes)	0.0027	(3.0)	0.0014	(1.7)	0.0111	(2.5)	0.0073	(1.5)
Broker Size (Log(Total Commiss))	0.0005	(2.6)	0.0002	(2.0)	0.0015	(1.7)	0.0013	(2.4)
Observations	7,597				7,597			
Funds	765				765			
Fund Family Fixed Effects	No		Yes		No		Yes	
R-Squared	63.0%		76.9%		17.7%		47.1%	

Table 5 – Performance, Expenses and Brokerage Commissions

Table 5 (next page) contains coefficient estimates and t-statistics from pooled regressions of fund performance on lagged fund characteristics. The sample includes all domestic equity funds with single-fund N-SAR filings over the period January 1996 through June 2009. The dependent variable is the fund's 6-month, forward-looking, 4-factor alpha [Carhart (1997)] using factor loadings estimated over the prior 36 months of data (t-1 to t-36). The independent variables include an intercept, the natural log of fund TNA and family TNA, turnover calculated as the minimum of fund purchases and sales divided by TNA and expense ratio (in specifications 1-3) and the components of the expense ratio (in specifications 4-7): distribution (12b-1) and non-distribution components (expense ratio minus 12b-1 fee). Specifications 2 through 6 include lagged total commissions, and lagged total commissions separated into a predicted and residual component using the model in specification 2 from Table 4. In specifications 4 through 6, we include an indicator variable for distribution support received by the fund from their broker and interact the same indicator variable with total commissions and residual total commissions. Specifications 5 and 6 also include an index fund indicator, the Herfindahl index of broker payment concentration, the percent of the fund in institutional share classes, a load and/or 12b-1 fund indicator, broker size and fund family fixed effects but these coefficients are not reported. Specification 7 replaces the predicted and residual total commissions (denominated by fund TNA), with predicted and residual commission rates (denominated by the sum of portfolio purchases and sales). These are interacted with the average trade volume (the sum of portfolio purchases and sales divided by TNA) of the investment objective. Also included is the ratio of the fund's trade volume to the average trade volume of the investment objective, interacted with the bundled distribution support indicator variable. The table reports p-values for two difference in coefficients tests and the coefficients being tested. Newey-West standard errors with 6 lags are used and the corresponding t-statistics are reported in parentheses.

Table 5 – Performance, Expenses and Brokerage Commissions (Continued)

Dependent Variable 6-Month Forward Looking 4-Factor Alpha

Regression:	1	2	3	4	5	6	7
Intercept	0.1406 (1.3)	0.1257 (-1.1)	0.0665 (0.6)	-0.1758 (-1.4)			-0.0912 (-0.7)
Log(Fund TNA)$_{t-1}$	-0.0212 (-3.4)	-0.0206 (-3.3)	-0.0190 (-2.9)	-0.0157 (-2.4)	-0.0256 (-3.3)	-0.0314 (-4.2)	-0.0170 (-2.7)
Log(Family TNA)$_{t-1}$	0.0092 (2.1)	0.0084 (1.9)	0.0084 (1.9)	0.0157 (3.5)	0.0010 (0.1)	-0.0017 (-0.2)	0.0136 (3.0)
Expense Ratio$_{t-1}$	-0.1052 (-0.6)	0.1943 (0.9)	0.3302 (1.6)				
Non-Dist. Exp. Rat.$_{t-1}$				0.9663 (3.3)	0.8821 (2.8)	0.7736 (2.5)	0.4711 (1.8)
Distribution Exp. Rat.$_{t-1}$				-1.3756 (-3.3)	-0.7570 (-1.2)	-0.6873 (-1.1)	-1.3292 (-3.1)
Turnover$_{t-1}$	-0.0002 (-1.0)	0.0002 (0.9)	0.0000 (0.2)	0.0000 (-0.0)	0.0000 (-0.2)	0.0001 (0.5)	
Total Commissions$_{t-1}$ (TC)		-1.8570 (-3.1)					
Pred. TC$_{t-1}$			-1.2064 (-1.4)	-1.2543 (-1.4)	1.2089 (1.1)		
Residual TC$_{t-1}$			-2.1425 (-3.1)				
Resid TC$_{t-1}$ x Broker Dist (=No)				-2.1211 (-2.6)	-1.0044 (-1.4)		
Resid TC$_{t-1}$ x Broker Dist (=Yes)				-4.1583 (-3.0)	-4.6058 (-3.6)		
Broker Distribution$_{t-1}$ (=Yes)				-0.0161 (-0.7)	-0.0167 (-0.5)	-0.1065 (-2.9)	-0.0895 (-2.8)
TC$_{t-1}$ x Broker Dist (=No)						0.2690 (0.4)	
TC$_{t-1}$ x Broker Dist (=Yes)						-4.8204 (-4.8)	
Pred. CR$_{t-1}$ x InvObjTradeVolume$_{t-1}$							-0.0476 (-0.3)
Resid CR$_{t-1}$ x InvObjTradeVolume$_{t-1}$ x Broker Dist (=No)							0.0674 (0.7)
Resid CR$_{t-1}$ x InvObjTradeVolume$_{t-1}$ x Broker Dist (=Yes)							-1.6315 (-3.0)
(TradeVolume$_{t-1}$/InvObjTradeVolume$_{t-1}$) x Broker Dist (=No)							0.0106 (0.7)
(TradeVolume$_{t-1}$/InvObjTradeVolume$_{t-1}$) x Broker Dist (=Yes)							-0.0904 (-4.1)
Observations	45,514	45,514	45,514	45,106	45,106	45,106	45,106
R-Squared	0.12%	0.28%	0.33%	0.51%	9.78%	9.81%	0.53%
Fund Family Fixed Effects	No	No	No	No	Yes	Yes	No
Other Controls	No	No	No	No	Yes	Yes	No
Difference in Coef Test #1	-	-	Exp. Rat. vs. ResidTC	Dist. Exp. vs. ResidTC x BrokDist(Yes)	Dist. Exp. vs. ResidTC x BrokDist(Yes)	Dist. Exp. vs. TC x BrokDist(Yes)	-
Difference p-Value	-	-	0.002	0.050	0.009	0.001	-
Difference in Coef Test #2	-	-	-	NonDistExp vs ResidTC x BrokDist(No)	NonDistExp vs ResidTC x BrokDist(No)	-	-
Difference p-Value	-	-	-	0.002	0.028	-	-

Table 6 – SEC Prohibition on Commission Bundling - Performance, Expenses and Brokerage Commissions

On December 13th, 2004, the SEC banned the use of brokerage commissions to finance distribution. Table 6 repeats the analysis of Table 5 for the time period surrounding the change (2003 to 2006) and for the subsample of funds that used bundled commissions to pay for distribution prior to the ban, but eliminated this practice after the ban. The analysis is run under two different assumptions regarding the date that the fund stopped their practice of bundling distribution costs with commissions. In specification 1 we assume that the fund ceased this practice on the date of the first N-SAR report that indicates they no longer bundle distribution (N-SAR Disclosed Date). In specification 2 we assume that all funds end this practice on the date by when the SEC requires compliance (SEC Compliance Date). The dependent variable is the fund's 6-month, forward-looking, 4-factor alpha [Carhart (1997)] using factor loadings estimated over the prior 36 months of data (t-1 to t-36). The independent variables include an intercept, the natural log of fund TNA and family TNA, turnover calculated as the minimum of fund purchases and sales divided by TNA, and the distribution (12b-1) and non-distribution components (expense ratio minus 12b-1 fee) of the expense ratio. Also included is an indicator variable for distribution support received by the fund from their broker and the interaction between the same indicator variable and residuals commissions. The table reports a p-value for two difference in coefficients tests and describes the coefficients being tested. Newey-West standard errors with 6 lags are used and the corresponding t-statistics are reported in parentheses.

	N-SAR Disclosed Date		SEC Compliance Date	
Regression:	1		2	
Intercept	-0.4043	(-1.6)	-0.3830	(-1.5)
Log(Fund TNA)$_{t-1}$	-0.0133	(-1.2)	-0.0115	(-1.0)
Log(Family TNA)$_{t-1}$	0.0228	(2.7)	0.0211	(2.5)
Non-Dist. Exp. Rat.$_{t-1}$	0.5377	(1.2)	0.3941	(0.9)
Distribution Exp. Rat.$_{t-1}$	-1.4877	(-2.3)	-1.3226	(-2.0)
Turnover$_{t-1}$	-0.0004	(-1.1)	-0.0004	(-1.1)
Pred. TC$_{t-1}$	0.0215	(0.0)	0.3580	(0.3)
Resid TC$_{t-1}$ x Broker Dist (=No)	0.9959	(0.5)	0.8809	(0.5)
Resid TC$_{t-1}$ x Broker Dist (=Yes)	-6.7770	(-4.1)	-8.0389	(-4.3)
Broker Distribution$_{t-1}$ (=Yes)	-0.0047	(-0.2)	-0.0620	(-2.0)
Observations	7,034		7,034	
R-Squared	3.22%		4.15%	
Difference in Coef. Test #1	Dist. Exp. vs. ResidTC x BrokDist(Yes)		Dist. Exp. vs. ResidTC x BrokDist(Yes)	
Difference #1 p-Value	0.002		0.001	
Difference in Coef. Test #2	ResidTC x BrokDist (Yes) vs. BrokDist (No)			
Difference #2 p-Value	0.008		0.001	

Table 7 – Fund Flows, Expenses and Brokerage Commissions

Table 7 presents coefficient estimates and t-statistics from pooled regressions of net fund flow on lagged fund characteristics. The sample includes all domestic equity funds with single-fund N-SAR filings over the period January 1996 through June 2009. The dependent variable is the annual fund flow as a percent of TNA. The independent variables include an intercept, the concurrent annual flow to the fund's investment objective (in %), lagged fund performance as measured by the fund's 4-factor alpha [Carhart (1997)] using factor loadings estimated over the prior 36 months of data (t-1 to t-36), the natural log of the lagged fund TNA and family TNA, lagged fund turnover, the lagged distribution (12b-1) and non-distribution components (expense ratio minus 12b-1 fee) of the expense ratio and an indicator variable . Specifications 2 through 4 include lagged total commissions, and lagged total commissions separated into a predicted and residual component using the model in specification 2 from Table 4. Specification 4 also includes an indicator variable for distribution support received by the fund from their broker and interact the same indicator variable with the residual total commissions variable. Newey-West standard errors with 12 lags are used and the corresponding t-statistics are reported in parentheses.

	Dependent Variable: Annual Fund Flow$_t$ (%)							
Regression:	1		2		3		4	
Intercept	69.98	(11.9)	69.90	(11.8)	73.66	(12.3)	75.42	(11.8)
Annual Inv. Obj. Flows$_t$ (%)	0.67	(23.9)	0.67	(23.8)	0.67	(23.8)	0.65	(23.2)
4-Factor Alpha$_{t-1,t-3}$ (%)	1.85	(20.5)	1.85	(20.5)	1.81	(20.1)	1.82	(20.2)
Log(Fund TNA)$_{t-1}$	-5.39	(-17.2)	-5.37	(-17.1)	-5.48	(-17.5)	-5.51	(-17.1)
Log(Family TNA)$_{t-1}$	1.53	(7.7)	1.55	(7.7)	1.51	(7.5)	1.42	(7.0)
Turnover$_{t-1}$ (%)	0.00	(0.8)	-0.01	(-1.2)	0.00	(0.3)	0.00	(0.0)
Non-Dist. Exp. Rat.$_{t-1}$	-4.02	(-4.4)	-4.73	(-5.1)	-4.55	(-4.8)	-4.04	(-3.8)
Distribution Exp. Rat.$_{t-1}$	3.04	(2.0)	2.51	(1.7)	2.87	(1.9)	3.11	(2.1)
Load Fund$_{t-1}$ (=Yes)	-3.65	(-4.3)	-3.56	(-4.2)	-3.45	(-4.1)	-4.86	(-5.6)
Total Commissions$_{t-1}$ (TC)			4.79	(3.6)				
Pred. TC$_{t-1}$					-3.49	(-1.7)	-3.74	(-1.8)
Residual TC$_{t-1}$					8.68	(4.3)		
Broker Distribution$_{t-1}$ (=Yes)							5.32	(6.5)
Resid. TC$_{t-1}$ x Broker Dist (=No)							7.96	(3.4)
Resid. TC$_{t-1}$ x Broker Dist (=Yes)							13.85	(3.2)
Observations	42,931		42,931		42,931		42,700	
R-Squared	21.52%		21.69%		22.00%		22.53%	

Table 8 – Multi-Fund Series Sample Characteristics

There are 179,798 monthly fund observations available from merging the Morningstar domestic equity fund data with the N-SAR database and applying our filters over the period January 1996 through June 2009. Table 1 presents descriptive statistics for the 45,514 observations with single-fund N-SAR filings. Table 8 presents descriptive statistics for the full sample of single and multi-fund N-SAR filing observations. The table reports the mean, median, 25th percentile and 75th percentile the annualized 4-factor alpha, fund TNA ($millions), family TNA ($billions), annual expense ratio, annual 12b-1 fees (for those funds that charge them – 109,134 observations), annual turnover (the minimum of fund purchases and sales divided by TNA in %), annual trade volume (fund purchases plus fund sales divided by TNA in %), a Herfindahl index of broker commissions (calculated using the dollar amount of commissions paid by each fund to their 10 highest compensated brokers, i.e., the maximum value of 1 means that all commissions were paid to a single broker and the minimum value of 0.1 means that commissions were evenly split among the 10 brokers), the percent of the fund's TNA in institutional share classes, broker size (broker size is the log of the annual total commissions received by the fund's primary broker), total commissions (annual brokerage commissions divided by TNA) and commission rates (brokerage commissions divided by the dollar value of purchases and sales over the same period). Panel B reports the number of observations by fund type, specifically the number of observations from index funds and broker-sold funds (i.e. funds that charge a load or a 12b-1 fee or both). Panel C reports the number of observations by Morningstar's investment objective and Panel D reports the number of observations and percent of the sample by the Morningstar style box.

Panel A. Univariate Statistics

Variable	Mean	Median	Percentiles 25th	75th
4-Factor Alpha (Annual)	-1.30%	-1.50%	-6.51%	3.51%
Fund TNA ($MM)	1,550.1	288.0	83.9	998.0
Family TNA ($BB)	55.1	9.1	1.4	44.0
Expense Ratio	1.35%	1.21%	0.91%	1.59%
12b-1 Fees (when > 0)	0.35%	0.28%	0.10%	0.51%
Turnover (%)	90%	62%	31%	111%
Trade Volume (% TNA)	205%	135%	71%	244%
Broker Commissions Herfindal	0.19	0.13	0.11	0.17
Institutional % of TNA	21.7%	0.0%	0.0%	27.0%
Broker Size (Annual Commiss in $MM)	166.5	59.9	9.5	356.2
Total Commissions (% TNA)	0.22%	0.13%	0.06%	0.26%
Commission Rate (% Trade Volume)	0.11%	0.09%	0.06%	0.13%

Panel B. Fund Type

	Obs.	Percent
Index Fund	10,791	6.0%
Broker Sold (Load/12b-1)	120,646	67.1%

Panel C. Investment Objectives

	Obs.	Percent
Aggressive Growth	8,018	4.5%
Growth	95,076	52.9%
Growth & Income	35,460	19.7%
Equity Income	8,589	4.8%
Small Company	32,655	18.2%

Panel D. Equity Style Box

	Value	Blend	Growth
Large Cap.	28,896	40,327	32,838
	16.1%	22.4%	18.3%
Mid Cap.	8,347	10,680	21,695
	4.6%	5.9%	12.1%
Small Cap.	8,989	11,214	16,812
	5.0%	6.2%	9.4%

Table 9 – Full Sample Brokerage Commission Model

Table 9 presents coefficient estimates and t-statistics from pooled regressions of brokerage commissions on fund characteristics. The sample includes all funds (from both single-fund and multi-fund series) with an N-SAR filing over the period January 1996 through June 2009. The observations are semi-annual (22,300 observations) to match the reporting of the N-SAR. The dependent variable is total commissions (brokerage commissions divided by TNA) in specifications 1 and 2 and commission rates (brokerage commissions divided by the dollar value of purchases and sales over the six month period) in specifications 3 and 4. Because the regression includes both single and multi-fund series, all fund level variables are aggregated to a series level. Fund TNA, for example, is replaced by Series TNA. Similarly, the index fund variable indicates the percent of the series that is in index funds. The independent variables include an intercept, domestic equity indicators (aggressive growth, growth, growth & income and income), as well as a bond, balanced and international indicators taken from the N-SAR filings. These investment objective variables are the percent of the series in each investment objective where the percentage is determined by fund TNA in the total commission regressions (specifications 1 and 2) and by trade volume in the commission rate regressions (specifications 3 and 4). The other independent variables include the natural log of series TNA and family TNA, series-level trading volume (fund purchases plus sales divided by fund TNA), an index indicator, a Herfindahl index of broker payment, the percent of the series in institutional share classes, a load and/or 12b-1 indicator, and broker size. Specifications (2 and 4) also include fixed effects for fund family. The standard errors are clustered by fund and the resulting t-statistics are provided in parentheses.

	Total Commissions (% TNA)				Commission Rates (% Trade Volume)			
Regression:	1		2		3		4	
Intercept	0.1116	(6.8)			0.3308	(17.0)		
Domestic Equity Invest. Style								
% of Series in:								
Agg. Growth	0.0108	(1.6)	0.0080	(1.7)	0.0246	(2.3)	0.0264	(3.4)
Growth	0.0088	(1.3)	0.0048	(1.0)	0.0228	(2.0)	0.0209	(2.7)
Growth & Income	0.0070	(1.1)	0.0035	(0.8)	0.0078	(0.7)	0.0107	(1.4)
Income	0.0046	(0.7)	0.0036	(0.8)	0.0082	(0.8)	0.0114	(1.5)
% of Series in Balanced	-0.0066	(-2.1)	-0.0083	(-4.0)	-0.0392	(-7.2)	-0.0405	(-9.1)
% of Series in Bond	-0.0191	(-3.2)	-0.0185	(-4.8)	-0.0636	(-7.3)	-0.0619	(-10.2)
% of Series in International	0.0108	(5.2)	0.0114	(5.8)	0.0639	(13.0)	0.0680	(13.4)
Log(Series TNA)	-0.0050	(-8.1)	-0.0054	(-9.7)	-0.0082	(-8.6)	-0.0079	(-8.8)
Log(Family TNA)	-0.0002	(-0.8)	-0.0003	(-1.1)	-0.0019	(-2.8)	-0.0014	(-2.5)
Trading Volume	0.0080	(12.1)	0.0074	(11.2)	-0.0028	(-9.2)	-0.0031	(-10.6)
% of Series in Index Funds	-0.0090	(-3.4)	-0.0067	(-2.4)	-0.0579	(-8.9)	-0.0520	(-6.4)
Broker Herfindahl	-0.0196	(-4.2)	-0.0206	(-5.6)	-0.0372	(-4.8)	-0.0369	(-5.7)
Institutional % of Series	-0.0026	(-0.9)	-0.0037	(-1.3)	-0.0117	(-2.2)	-0.0071	(-1.4)
% of Series charging Load/12b1	-0.0006	(-0.4)	-0.0004	(-0.2)	0.0019	(0.6)	0.0071	(2.1)
Broker Size (Log(Total Commiss))	0.0005	(2.1)	0.0006	(3.1)	-0.0008	(-1.6)	-0.0002	(-0.5)
Observations	22,300				22,300			
Funds	1,345				1,345			
Fund Family Fixed Effects	No		Yes		No		Yes	
R-Squared	52.8%		66.9%		26.3%		47.3%	

Table 10 – Full Sample Performance, Expenses and Brokerage Commissions

Table 10 contains coefficient estimates and t-statistics from pooled regressions of fund performance on lagged fund characteristics. The sample includes all domestic equity funds over the period January 1996 through June 2009. The dependent variable is the fund's 6-month, forward-looking, 4-factor alpha [Carhart (1997)] using factor loadings estimated over the prior 36 months of data (t-1 to t-36). The dependent variable is the fund's 6-month, forward-looking, 4-factor alpha [Carhart (1997)] using factor loadings estimated over the prior 36 months of data (t-1 to t-36). The independent variables include an intercept, the natural log of fund TNA and family TNA, turnover calculated as the minimum of fund purchases and sales divided by TNA and expense ratio (in specifications 1-3) and the components of the expense ratio (in specification 4): distribution (12b-1) and non-distribution components (expense ratio minus 12b-1 fee). Specifications 2 through 4 include lagged total commissions, and lagged total commissions separated into a predicted and residual component using the model in specification 4 from Table 9. In specification 4, we include an indicator variable for distribution support received by the fund from their broker and interact the same indicator variable with total commissions and residual total commissions. The table reports p-values for two difference in coefficients tests and the coefficients being tested. Newey-West standard errors with 6 lags are used and the corresponding t-statistics are reported in parentheses.

	Dependent Variable 6-Month Forward Looking 4-Factor Alpha							
Regression:	1		2		3		4	
Intercept	0.1743	(3.4)	0.1813	(3.5)	0.1785	(3.5)	0.0955	(1.8)
Log(Fund TNA)$_{t-1}$	-0.0236	(-8.5)	-0.0243	(-8.7)	-0.0228	(-8.1)	-0.0222	(-7.7)
Log(Family TNA)$_{t-1}$	0.0089	(4.2)	0.0091	(4.3)	0.0081	(3.8)	0.0110	(5.0)
Expense Ratio$_{t-1}$	-0.2693	(-4.0)	-0.1827	(-2.5)	-0.2691	(-3.6)		
Non-Dist. Exp. Rat.$_{t-1}$							0.0225	(0.2)
Distribution Exp. Rat.$_{t-1}$							-1.2047	(-6.9)
Turnover$_{t-1}$	0.0001	(0.9)	0.0001	(1.2)	0.0000	(0.6)	0.0001	(0.8)
Total Commissions$_{t-1}$ (TC)			-0.5302	(-2.2)				
Pred. TC$_{t-1}$					-0.2996	(-1.3)	-0.3314	(-1.4)
Residual TC$_{t-1}$					-1.5538	(-4.3)		
Resid TC$_{t-1}$ x Broker Dist (=No)							-0.9701	(-2.5)
Resid TC$_{t-1}$ x Broker Dist (=Yes)							-4.0419	(-5.2)
Broker Distribution$_{t-1}$ (=Yes)							-0.0061	(-0.6)
Observations	179,798		179,798		179,798		174,581	
R-Squared	0.20%		0.23%		0.27%		0.31%	
Difference in Coef Test #1	-		-		Exp. Rat. vs. ResidTC		DistExp vs. ResidTC x BrokDist(Yes)	
Difference p-Value	-		-		0.001		0.001	
Difference in Coef Test #2	-		-		-		NonDistExp vs ResidTC x BrokDist(No)	
Difference p-Value	-		-		-		0.019	

Table 11 – Full Sample SEC Prohibition on Commission Bundling - Performance, Expenses and Brokerage Commissions

On December 13th, 2004, the SEC banned the use of brokerage commissions to finance distribution. Table 11 repeats the analysis of Table 10 for the time period surrounding the change (2003 to 2006) and for the subsample of funds that used bundled commissions to pay for distribution prior to the ban, but eliminated this practice after the ban. The analysis is run under two different assumptions regarding the date that the fund stopped their practice of bundling distribution costs with commissions. In specification 1 we assume that the fund ceased this practice on the date of the first N-SAR report that indicates they no longer bundle distribution (N-SAR Disclosed Date). In specification 2 we assume that all funds end this practice on the date by when the SEC requires compliance (SEC Compliance Date). The dependent variable is the fund's 6-month, forward-looking, 4-factor alpha [Carhart (1997)] using factor loadings estimated over the prior 36 months of data (t-1 to t-36). The independent variables include an intercept, the natural log of fund TNA and family TNA, turnover calculated as the minimum of fund purchases and sales divided by TNA, and the distribution (12b-1) and non-distribution components (expense ratio minus 12b-1 fee) of the expense ratio. Also included is an indicator variable for distribution support received by the fund from their broker and the interaction between the same indicator variable and residuals commissions. The table reports a p-value for two difference in coefficients tests and describes the coefficients being tested. Newey-West standard errors with 6 lags are used and the corresponding t-statistics are reported in parentheses.

	N-SAR Disclosed Date		SEC Compliance Date	
Regression:	1		2	
Intercept	-0.4162	(-3.6)	-0.4114	(-3.6)
Log(Fund TNA)$_{t-1}$	-0.0087	(-1.6)	-0.0084	(-1.6)
Log(Family TNA)$_{t-1}$	0.0205	(4.4)	0.0204	(4.4)
Non-Dist. Exp. Rat.$_{t-1}$	0.4269	(2.5)	0.3933	(2.3)
Distribution Exp. Rat.$_{t-1}$	-1.9354	(-6.8)	-1.9027	(-6.8)
Turnover$_{t-1}$	-0.0006	(-3.7)	-0.0005	(-3.4)
Pred. TC$_{t-1}$	0.2009	(0.5)	0.2566	(0.7)
Resid TC$_{t-1}$ x Broker Dist (=No)	0.7946	(1.0)	0.2378	(0.3)
Resid TC$_{t-1}$ x Broker Dist (=Yes)	-5.3173	(-6.9)	-5.4415	(-6.2)
Broker Distribution$_{t-1}$ (=Yes)	-0.0443	(-3.2)	-0.0976	(-6.4)
Observations	21,540		21,540	
R-Squared	1.92%		2.37%	
Difference in Coef. Test #1	Dist. Exp. vs.		Dist. Exp. vs.	
Difference #1 p-Value	0.001		0.001	
Difference in Coef. Test #2	ResidTC x BrokDist (Yes) vs. BrokDist (No)			
Difference #2 p-Value	0.001		0.001	

33

Brokers vs. Retail Investors:

Conflicting Interests and Dominated Products

Mark Egan*

November 11, 2014

Abstract

I study how brokers distort consumer investment decisions. The market for retail convertible bonds offers a unique environment to study consumer investment decisions in a broker-intermediated setting. Using a novel data set, I find that consumers frequently purchase dominated bonds in this market – i.e., cheap and expensive versions of otherwise identical bonds exist in the market at the same time. Moreover, inconsistent with standard search models, consumers purchase more of the expensive bonds. The empirical evidence suggests broker incentives are partially responsible for the inferior investments as brokers earn a 1.12% point higher fee for selling the dominated bond. I rationalize the behavior of brokers and consumers in equilibrium by developing and estimating a search model. Consumer search is endogenously directed according to the incentives of brokers and a broker's ability to price discriminate across consumers based on the consumer's level of sophistication. I use the estimated model to disentangle and quantify the importance of search, consumer sophistication, and broker incentives. Aligning broker incentives with those of consumers' increases consumer risk-adjusted returns by 80bps, but does not resolve the primary friction in this market, consumer search. My estimated model allows for investigation of counterfactual scenarios surrounding the Dodd-Frank Act.

*Mark Egan: University of Chicago Department of Economics, 1126 East 59th Street Chicago, IL 60637, egan@uchicago.edu. I would like to thank my advsiors Ali Hortaçsu, Gregor Matvos, Tobias Moskowitz and Amit Seru for their help and support. I am grateful for the comments received from Jesse Shapiro, Eric Budish, Juhani Linnainmaa, Neale Mahoney, Monika Piazzesi, Chad Syverson, Derek Neal, Tomas Philipson, Casey Mulligan, Frank Wolak, Amir Sufi and the members of the University of Chicago Booth School of Business Fama Miller Corporate Finance Reading Group as well as seminar participants at the North American Econometric Society Meetings, the University of Chicago Booth School of Business Finance Seminar, the University of Chicago Workshop in the Applications of Economics, the Washington University in St. Louis Graduate Student Conference, and AQR Capital Management.

1 Introduction

The prices and fees of seemingly identical financial products often differ drastically. Previous research documents price heterogeneity across mutual funds, mortgages, bonds and other financial products.[1] Does the observed price dispersion imply that some consumers are overpaying for investment opportunities? If so, what is driving this behavior? Sirri and Tufano (1998) and Hortaçsu and Syverson (2004) highlight the importance of search in a consumer's investment decision process. However, consumer search does not happen in a vacuum. Broker intermediation plays a critical role in a consumer's investment decision and search process. In 2010, 56% of American households sought investment advice from a financial professional.[2] Despite their prevalence, brokers may not be acting in the best interests of their clients. A broker may choose to subordinate her client's interests for her own financial interests by directing her client to inferior products with high broker's fees. While arguments such as these are abundantly available and have guided much of the policy response in the aftermath of the crisis (see section Section 913 of the Dodd-Frank Act), a rigorous empirical and theoretical investigation of this issue has been lacking. In this paper I fill this gap.

The paper has two goals. The first goal is to use novel data and a unique setting to show that consumers frequently purchase the dominated product in a market – i.e., cheap and expensive versions of otherwise identical products exist in the market at the same time – and that broker incentives are partially responsible for the inferior investments. The second goal is to rationalize the behavior of brokers and consumers in equilibrium by developing and estimating a search model. The model helps disentangle and quantify the importance of search, consumer sophistication, and broker incentives. I also use the model to investigate counterfactual scenarios surrounding the Dodd-Frank Act.

Several data challenges contribute to the lack of studies investigating theses issue in detail. First, with most financial products, it is hard to find scenarios where one can easily compare products and rank one product as unambiguously dominating the other. Financial products, such as mutual funds, differ on a plethora of observable and unobservable characteristics, making direct comparisons

[1]For examples, see Hortaçsu and Syverson (2004) and Elton et al. (2004) for the mutual fund industry, Gurun, Matvos and Seru (2013) for mortgages, Green Hollifeld and Shürhoff (2007) for bonds, Duarte and Hastings (2012) for privatized social security plans, Christoffersen and Musto (2002) for money funds, and Brown and Goolsbee (2002) for life insurance.

[2]Source: Survey of Consumer Finances

of products tenuous. We may think that a consumer paying 2% for an S&P index fund is overpaying for that investment product. However, without observing all of the fund characteristics, making such claims is impossible. The problem is compounded once we allow for heterogeneity across consumer preferences and portfolio holdings. Some consumers may inherently prefer Vanguard funds to Fidelity funds or vice versa. Alternatively, different retirement plans may restrict the set of fund families consumers are eligible to hold. Second, little data have been available on the compensation of financial intermediaries. Did a consumer buy mutual fund XYZ or was he sold mutual fund XYZ by his broker?

I address these challenges by constructing a new retail bond data set covering reverse convertible bonds issued in the United States over the period 2008-2012. A reverse convertible is a fixed rate bond for which the final principal payment is convertible into shares of some pre-specified equity. The advantage of studying reverse convertible bonds over other financial products is twofold. First, reverse convertibles are completely characterized by a small number of dimensions, namely, a fixed coupon and an equity-linked principal payment. As a result, simultaneously issued reverse convertibles for which the payout of one reverse convertible is unambiguously dominated by another – the bond with the higher coupon – are easy to locate. Consider the following two nearly identical one-year reverse convertibles issued by JPMorgan Chase on June 30, 2008.[3] One reverse convertible pays a fixed coupon of 11.25%, whereas the other pays a fixed coupon of 9.00%. Both reverse convertibles were sold to investors at a fixed par price of 100%. The final principal payment of both reverse convertibles is identical and linked to the share price of Microsoft Inc. If the price of Microsoft Inc. shares ever closes below $22.68, the bond principal (for both bonds) is converted into equity where bond holders receive at maturity 35.27 shares of Microsoft Inc. for every $1,000 invested.[4] Figure 1 displays the hypothetical return to investors of the two products. Notice that the return of the 11.25% reverse convertible clearly dominates that of the 9.00%.[5] However, in practice, consumers purchased more than 10 times as much of the dominated product. This example of a bank simultaneously issuing a dominated/superior product is not unique; I observe over 100 dominated/superior reverse convertibles in the data set.

[3]CUSIPS: 48123LAM6 and 48123LBR4

[4]The principal payment on both reverse convertibles is capped at par.

[5]The example of unambiguously dominated structured products is interesting when contrasted with the work of Carlin (2009) and Célérier and Vallée (2014). Banks could easily make the differences across products less salient by changing either the convertible price or the underlying equity; however, they often choose not to.

Figure 1: Dominated Reverse Convertible Example

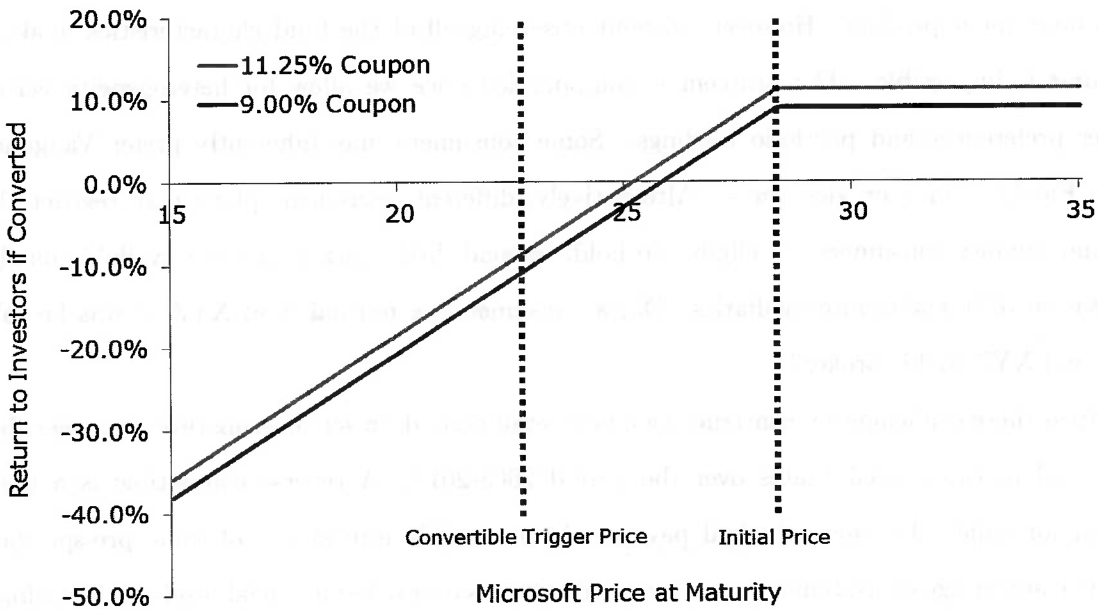


Figure 1 Notes: The figure displays the return to investors for two one-year reverse convertible bonds linked to the price of Microsoft Inc. that were issued by JPMorgan Chase on June 30th, 2008. The reverse convertibles pay monthly coupons of 9.00% and 11.25%, respectively. If the share price of Microsoft Inc. ever closes below the convertible trigger price of $22.68 during the life of the reverse convertible, the issuer will pay the bondholder 35.27 shares of Microsoft Inc. per $1,000 invested ($1,000/Initial Price) rather than 100% of the principal amount invested. The final principal payment paid by the issuer is capped at par (100%). The above figure displays the final return to investors provided the principal has been converted into equity. Note that the 11.25% convertible always yields a 2.25% higher return than the 9.00% convertible.

The second advantage of studying reverse convertibles is that the Securities and Exchange Commission (SEC) requires all bond issuers to disclose the fees/commissions paid to brokers. Reverse convertible bond issuers, rather than consumers, compensate brokers with fees for selling reverse convertibles. In the previous JPMorgan example, JPMorgan paid brokers a commission of 3.09% for selling the worse 9.00% reverse convertible and only 2.15% for selling the better reverse convertible. This data therefore allows me to quantify the degree to which broker incentives influence consumer choice.

Using this new retail bond data set, I analyze the investment decisions of consumers in a broker-intermediated market. I find clear evidence of consumers buying unambiguously dominated financial products. Simultaneously issuing identical retail bonds at the same price with different interest rates/coupons such that the payout of one bond unambiguously dominates that of the other is a common practice for investment banks. By simply buying the superior product, consumer risk-adjusted returns would have increased by 1.60% on average. What is more staggering, is that when both a dominated and superior product were available, consumers collectively purchased 16% more of the dominated product. The prevalence of dominated products suggests consumers purchasing the dominated product are not aware of the superior product. Hence, a consumer's investment decision problem is fundamentally a search problem. Search helps explain why consumers buy dominated products, but a standard search model[6] predicts that consumers would purchase more of the better product, which runs contrary to what is observed in the data. I argue that financial product distribution rationalizes such behavior. The incentives of brokers do not always align with the incentives of consumers. Evidence from the retail bond data set indicates that, all else equal, consumers are more likely to buy products with higher brokerage fees, and products with higher brokerage fees have worse payoff profiles. On average, brokers earned a 1.12% point higher fee for selling dominated products.

The first part of my paper reveals three stylized empirical facts. First, the risk-adjusted returns of reverse convertibles exhibit substantial dispersion, and consumers often fail to purchase the best available financial product. The standard deviation of risk-adjusted returns in the data set is over 2.40%. Second, when better and worse products are available, consumers actually purchase more of the worse product. Third, the evidence suggests the incentives of brokers do not align with the incentives of consumers. All else equal, consumers collectively tend to purchase more products with higher fees, and products with higher fees have lower payoffs.[7] I argue that the incentive and information asymmetry between brokers and consumers helps rationalize product issuance and the behavior of consumers.

In the second part of the paper, I rationalize the behavior of brokers and consumers in equilibrium by developing and estimating a search model. The model helps disentangle and quantify the importance of search, consumer sophistication, and broker incentives. In the model, consumers sequentially search for investment products with the aid of a broker. Brokers service their customer base by offering different products to each client. Brokers select products to offer each client based

[6]A sequential search model with random undirected search would predict that consumers purchase more of the better product.

[7]The finding that high fee products have worse risk-adjusted returns is consistent with evidence Gil-Bazo and Ruiz-Verdú (2009) find in the mutual fund industry.

on the quality of the financial product and the underlying broker's fee. In other words, broker profit maximization endogenously determines the distribution of products that consumers observe. Consumers ultimately decide whether to purchase the offered product or continue searching. Consumers differ in their level of financial sophistication (measured as search costs), and brokers utilize the full product space to price discriminate across consumers based on the consumer's level of financial sophistication.

The model introduces two frictions that are consistent with the empirical data. First, consumers must engage in costly search for products which explains why consumers might purchase inferior products. Second, brokers are incentivized to show high-fee products, which makes finding better products relative to worse products potentially harder for consumers. Broker's incentives in conjunction with consumer search help explain why consumers generally fail to purchase the best available products. I structurally estimate this model using the reverse convertible data set to determine whether the frictions in the model and associated costs are economically meaningful.

The model provides sharp insights that are useful in understanding consumer and broker behavior beyond just the reverse convertible market. First, the model helps to evaluate if the search costs and broker behavior that help rationalize the empirical facts documented earlier are "reasonable". Second, I can assess the total cost of each friction. For example, the model estimates suggest the average consumer spends over $207 (in terms of the opportunity cost of time and the cost of delaying investment) searching for a $10,000 investment. Third, I am able to show that aligning broker incentives with those of consumers' would increase consumer risk-adjusted returns increase by 80bps, but does not resolve the primary friction in this market, consumer search. Aligning the incentives of brokers helps consumers search more effectively. However, consumers still have to engage in costly search. This result speaks directly to policies passed as a part of the Dodd-Frank Act where the regulators may soon hold brokers to a fiduciary duty. Holding brokers to a fiduciary duty would force brokers to act in the best interest of their clients, which could result in consumers holding better financial products. The model suggests that while this would alleviate the consumer search problem, it would not eliminate it. Finally, my estimated model allows for investigation of other counterfactual scenarios surrounding the Dodd-Frank Act and issues related to optimal financial regulation in general.

This paper relates to the economics and finance literature regarding price and quality dispersion in financial products. Previous work including but not limited to Massa (2000), Hortaçsu and Syverson (2004), Choi et al. (2010), Wahal and Wang (2011), and Khoran and Servaes (2012) indicate the law of one price may fail to hold in the mutual fund industry. Similarly, Anagol et al. (2012) find similar evidence in life insurance markets in India. One limitation of previous studies is that much of the observed dispersion in prices and quality of financial products could potentially be rationalized by unobserved product characteristics and preference heterogeneity. This paper offers the cleanest setting for studying retail financial markets. All consumers would be unambiguously better off purchasing the superior reverse convertible over the dominated convertible regardless of the consumer's preferences or portfolio.

Researchers have documented the potential broker and consumer information and incentive asymmetry arising in consumer finance (Livingston and O'Neal 1996, Mahoney 2004, Bolton et al. 2007, Bergstresser et al. 2009, Woodward and Hall 2012, Christoffersen et al. 2013). I find evidence consistent with Bergstresser et al. (2009), Anagol et al. (2012), and Christoffersen et al. (2013) suggesting that brokers may direct consumers into high-fee products. This paper builds on the preceding work by studying financial distribution in a clean setting in which identifying the conflict-of-interest problem is easier. In the data set, I observe all product characteristics as well as the fees paid to brokers. By directly comparing the dominated and superior products, I can isolate the effect of broker's fees on product issuance. The previous research suggests that underlying economic frictions in the market for reverse convertibles, search and broker incentives, apply to a much broader set of financial markets.

The remainder of the paper is laid out as follows. In Sections 2 and 3, I describe the reverse convertible data set and some fundamental features of the reverse convertible market. In Section 4, I analyze the reverse convertible data set and examine the characteristics of reverse convertibles purchased by consumers. In Sections 5 and 6, I develop and then structurally estimate a search model of financial distribution. I report the corresponding structural estimation results in Section 7. In Section 8, I use the structural estimates to quantify the inefficiencies in retail financial markets and evaluate the proposed regulatory response. Lastly, Section 9 concludes the paper.

2 Institutional Background: Reverse Convertibles

The empirical analysis focuses on the market for lightly structured retail bonds, specifically equity reverse convertibles. A standard fixed-rate bond consists of a set of fixed coupon payments and final principal payment at maturity. Reverse convertible securities are similar to fixed-rate bonds except the final principal payment can be converted into shares of equity. At maturity, investors receive 100% of their principal provided that the underlying equity remains above the pre-specified convertible trigger price. If the equity falls below the convertible trigger price during the life of the bond, investors receive a fixed number of equity shares rather than the full principal amount.[8] The value of the shares may be worth substantially less than the initial principal amount invested.

A reverse convertible essentially combines a standard fixed rate bond and an equity put option into one financial product. By buying a reverse convertible, the bondholder effectively sells the issuer a knock-in European put option. As illustrated in Figure 1, the bondholder is short a Microsoft Inc. knock-in put option that is struck at the initial share price of $28.35 and knocks-in at the convertible trigger price of $22.68. The issuer uses the premium earned from the knock-in put option to fund the broker's fee and the coupon paid to the bondholder.

2.1 The Market for Reverse Convertibles

Reverse convertibles offer a unique setting for understanding consumer investments and studying retail financial distribution. The financial industry largely recognizes reverse convertibles as the "Gold Standard" of retail structured products. Banks issued almost $5 billion of reverse convertibles in the US in 2011 and $50 billion globally, the bulk of which were purchased by retail investors.[9] Reverse convertibles are largely an access product, allowing purchasers to sell equity options/volatility, which makes these products desirable for retail consumers rather than for companies and professional investors. Reverse convertibles provide investors with an opportunity to enhance the yield on a standard three month to two year fixed rate bond by taking some additional equity risk. Reverse convertibles pay a fixed, guaranteed, relatively high[10] interest rate over the

[8] In practice two different common types of reverse convertibles exist: single observation and continuous observation. The previous discussion describes a continuous observation reverse convertible. The single versus continuous observation reverse convertibles differ with respect to the principal payment at maturity. A single observation reverse convertible is converted into equity if the equity price is below the convertible trigger price at maturity rather than if the equity price is ever below the convertible trigger price. Figure A-1 in the appendix walks through an example of a single observation reverse convertible.

[9] Source: Bloomberg

[10] The interest payments for the average reverse convertible in the sample exceeded ten percent per annum; the interest rate for a corresponding fixed rate bond was less than two percent over the same period.

life of the bond. To protect retail consumers, the SEC requires disclosure of the details of each reverse convertible issued, including broker's fees. Though relatively simple, reverse convertibles are often synonymous with structured products which are often criticized for their opaqueness and high costs.[11] The complexity and prevalence of reverse convertibles makes them of particular importance when analyzing some of the new proposed SEC broker regulations.

One of the primary advantages of studying reverse convertibles is that that they are relatively easy to compare and contrast. Reverse convertibles are completely characterized by a small number of observable dimensions. A reverse convertible consists of an issuer, fixed coupon, broker's fee and equity put option. Additionally, it is common practice for banks to issue reverse convertibles that are unambiguously dominated. Banks frequently issue two reverse convertibles with the exact same risk and payout profiles; however, one reverse convertible will have a relatively high fixed coupon and a low broker's fee while the other has a relatively low fixed coupon and a high broker's fee. By studying the purely dominated/superior reverse convertibles, I am able to measure how consumers and brokers trade-off coupon and fees while controlling for all other product characteristics.

2.2 Reverse Convertible Market Structure and Distribution

The reverse convertible market consists of three players: product issuers, brokers and retail consumers. Product issuers, banks, create and issue reverse convertibles. Brokers purchase reverse convertible bonds from the product issuer and then sell the bonds to retail consumers..

Figure 2 illustrates the reverse convertible distribution process. Typically, at the beginning of each month product issuers create a suite of available reverse convertibles that will be issued at the end of the month. The issuer fixes all of the characteristics of each reverse convertible, including the broker's fee, at the beginning of each month.[12] Over the course of the month, issuers market available reverse convertibles to brokers who then solicit orders from retail consumers. At the end of the month all of the orders are accumulated and the reverse convertible is issued such that demand is completely satisfied. Issuers sell the reverse convertibles at a fixed par price of 100% minus the fixed broker's fee. Brokers then sell reverse convertibles to the end consumer at a fixed price of

[11]See Stoimenov and Wilkens (2005), Henderson and Pearson (2011) and Szymanowska et al. (2009) for further details.

[12]Since the initial equity price is not known prior to issuance, the convertible trigger price is fixed and expressed as a percentage of the initial equity price.

par (100%).[13] For each product sold, brokers earn the broker's fee. Since issuers pay the fee, it represents a transfer from the issuer to the broker. Consequently consumers are ambivalent over the broker's fee conditional on the risk and return of the product.

FIGURE 2: REVERSE CONVERTIBLE DISTRIBUTION

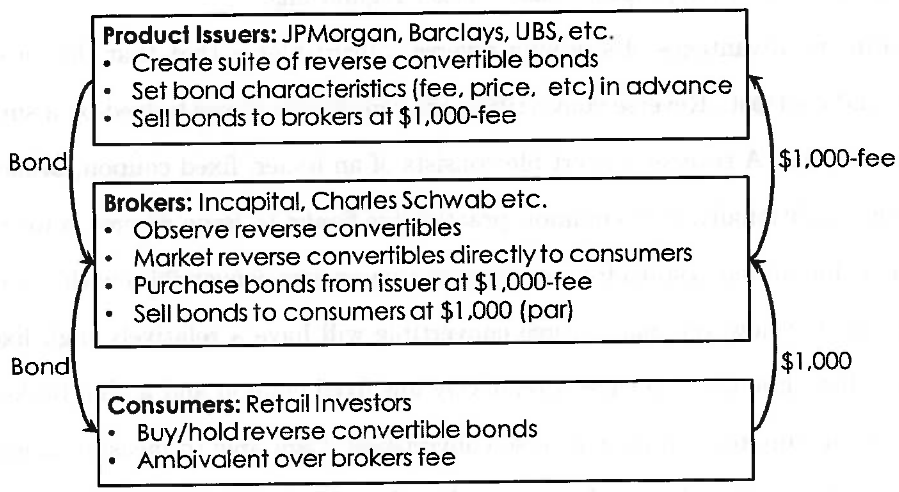


Figure 2 Notes: The figure displays the market structure/ distribution process of the reverse convertible market.

For regulatory reasons, issuers sell reverse convertibles through brokerage houses rather than selling them directly to consumers. SEC regulations such as the Securities Act of 1933 restrict the marketing of financial products to end consumers. Any materials used to market an SEC registered security (such as the reverse convertibles studied here) must be vetted for legal and compliance reasons and formally filed with the SEC. Since creating marketing materials can be a costly and lengthy process relative to the marketing period (typically one month), issuers do not market reverse convertibles directly to consumers.[14] Rather, issuers choose to sell reverse convertibles to brokers who market them to consumers directly.

3 Data and Summary Statistics

The empirical analysis uses a new reverse convertible bond data set constructed for this paper.

[13]The majority of reverse convertibles are fixed price par offerings which means that they must be sold at a fixed price of par. On occasion, certain banks will issue reverse convertibles as variable price re-offerings which means they could theoretically be sold at a discount.

[14]Previous research such as Jain and Wu (2000), Cronqvist and Thaler (2004), Barber et al. (2005), Cronqvist (2006), and Hastings et al. (2013) find that advertising plays a critical role in the competition and demand for financial products.

The data set covers US, SEC registered, one year maturity reverse convertibles issued over the period 2008-2012. Issuance data, specifically the date, coupon, and size details are from Bloomberg and the Mergent Fixed Income Securities Database data sources. Details on each reverse convertible's broker's fees, initial equity share price and convertible trigger price were manually collected from the corresponding Form 424b filings found on the SEC EDGARS website. The data set is supplemented with equity volatility data from Option Metrics and Credit Default Swap (CDS) data from Markit.

Table 1 displays the summary statistics of the data set. The mean and median issuance size in the sample was \$1.64 million and \$665 thousand respectively. To ensure that the data set is limited to retail consumers, the largest 1% issuances (exceeding \$17.51 million) are dropped from the data set. On average, reverse convertibles paid a coupon of 10.50% per annum. The option premium measures the value of the put options embedded in each reverse convertible expressed as a percentage of the notional invested.[15] The one year credit default swap (CDS) spread reflects the default risk for senior unsecured debt which corresponds to the issuer credit risk inherent to each reverse convertible bond.

TABLE 1: SUMMARY STATISTICS

Variable	Obs	Mean	Std.Dev.	Min	Max
Size (mm)	3066	1.64	2.62	0.00	17.51
Coupon	3066	10.50%	2.84%	3.24%	27.00%
Option Premium	3066	16.10%	3.90%	2.55%	42.90%
Fee	3066	2.24%	0.70%	0.00	6.75%
CDS Spread	2680	0.78%	0.60%	0.04%	9.20%

Table 1 Notes: Table 1 reflects US SEC registered one year equity reverse convertible issuance data over the period 2008-2012.

Reverse convertibles are almost exclusively issued by banks. Five banks: ABN Amro, Barclays Bank, JPMorgan Chase & Co, UBS and Royal Bank of Canada, dominate the issuance market for

[15]Option prices were calculated according to the Black Scholes (1973) formula for standard European options and the Reiner and Rubinstein (1991a 1991b) formulas for knock-in options. For a summary of the formulas see Haug (2007). I assume each underlying equity pays a constant dividend. The implied dividends are backed out from Option Metrics option price data.

one year reverse convertibles, making up over 80% of the market over the period 2008-2012. Apple Inc. served as the most popular underlying equity to link reverse convertibles to. Other popular underlying equities include Bank of America Corporation, General Electric Company, Caterpillar Inc., and JPMorgan Chase & Co.

4 What Type of Reverse Convertible Bonds Do Consumers Buy?

In this section, I examine the characteristics of reverse convertibles purchased by consumers. As alluded to in the introduction, consumers do not always purchase the best available convertibles. Using the new reverse convertible data set, I first examine the dispersion in reverse convertible returns. Second, I look at how often do consumers purchase inferior products. And third, I investigate how the incentives of brokers might drive this behavior. I use the prevalence of dominated products to isolate key variation in product returns and fees, which makes this a clean setting to study financial product distribution.

4.1 Dispersion in Reverse Convertible Returns

An extensive economics literature documents the substantial heterogeneity in the fees and prices of seemingly identical financial products. I find similar and perhaps the cleanest evidence of the failure of the law of one price in the market for reverse convertible bonds. Note that since because reverse convertibles are issued at a fixed price of 100%, I do not observe price dispersion per se, but rather dispersion in the potential returns to investors. If two otherwise identical bonds are being sold with different coupons, this event is analogous to the same product being sold at different prices.

I calculate the dispersion in reverse convertible coupons conditional on all other observable bond characteristics. I examine the variation in reverse convertible coupons conditional on the value of the embedded put option and the issuer's CDS spread. Specifically, I estimate the linear regression

$$Coupon_j = \beta_0 + \beta_1 Option\_Premium_j + \beta_2 CDS_j + Fixed\_Effects \qquad (1)$$

where I also include issuer, month, and equity fixed effects and control for the type of reverse convertible issued (single vs. continuous observation). Observations are reverse convertible bond issuances. Theoretically, the above specification should control for all relevant reverse convertible characteristics. If the law of one price holds, we would expect to see little variation in coupons conditional on all other product characteristics.

Figure 3 plots the residuals from regression (1) and reflects dispersion in reverse convertible coupons conditional on observable product characteristics. Provided that specification (1) correctly captures all product characteristics, the observed dispersion in coupons is analogous to dispersion in investor risk-adjusted returns. Conditional on product characteristics, the reverse convertible market exhibits substantial heterogeneity in coupons. The standard deviation of conditional coupons is 1.14%. The results suggest the investor purchasing the best reverse convertible would earn a return that is over 10% higher than the investor purchasing the worst reverse convertible on a risk-adjusted basis. Relative to the average risk-free rate over the period studied, 0.60%, the dispersion in coupons is substantial. As a robustness check, I calculate return dispersion in terms of the risk neutral value of each reverse convertible and find further evidence supporting the heterogeneity in reverse convertible returns (see Figure A-2 in the appendix). The standard deviation of risk-adjusted returns is over 2.40%.

FIGURE 3: DISPERSION IN REVERSE CONVERTIBLE RETURNS

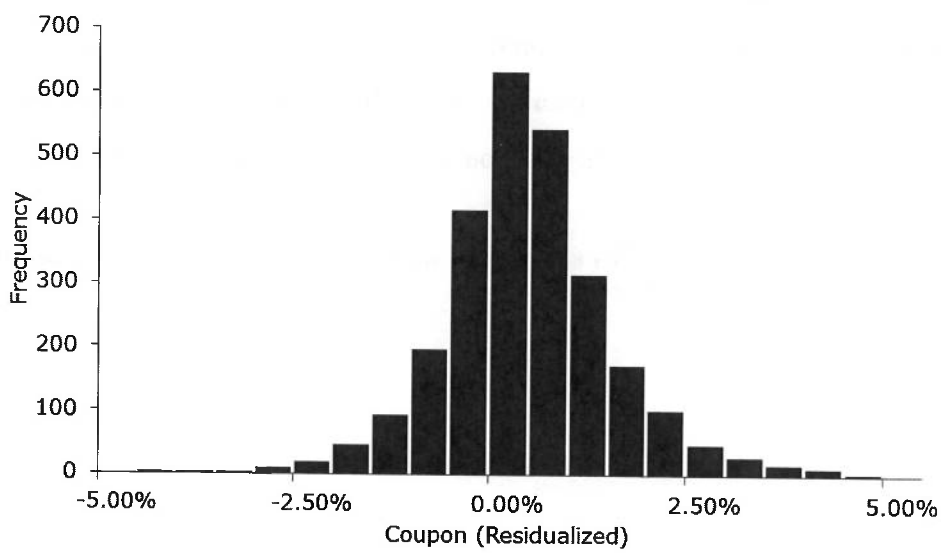


Figure 3 Notes: The figure displays dispersion in reverse convertible coupons. Specifically, Figure 3 plots the residual from the regression of reverse convertible coupons on all observable product characteristics.

The dispersion observed in Figure 3 stems from two potential sources. First, the observed dispersion may be "real" coupon dispersion such that the conditional returns of some reverse convertibles are simply higher than others. Second, the dispersion could be a function of unobserved product characteristics. In the case of mutual funds, previous research finds that unobserved fund characteristics are economically meaningful (Hortaçsu and Syverson 2004). The presence of unobserved product characteristics is likely less of a concern with reverse convertibles as their payouts are fully characterized by a small number of observable dimensions. To isolate "real" coupon dispersion from unobserved product heterogeneity, I examine the set of reverse convertibles that are either unambiguously dominant or are dominated by other reverse convertibles.

As discussed in the introduction, banks frequently issue identical reverse convertibles at the same price with two different coupon rates. I define a reverse convertible as being dominated if another reverse convertible exists with the same issuer, convertible payout, issue date, and price, with a higher coupon rate. In the data set of 3,066 reverse convertible bonds, 142 either dominate or are dominated by another reverse convertible.

Figure 4 plots coupon dispersion within dominated/superior products. I define a set of dominated/superior products as all products with the same issue date, issuer, price, and convertible payout. Figure 4 plots the distribution of $c_{i,j} - \bar{c}_j$. Here, i indexes the reverse convertible issuance and j indexes a set of dominated/superior products. By plotting $c_{i,j}$ relative to $\bar{c}_j$, I am able to perfectly control for all other product characteristics. Hence, Figure 4 plots the "real" coupon dispersion.[16] Figure 4 also reflects the distribution of ex-ante and ex-post returns among reverse convertibles.

[16]Another way of constructing Figure 4 would be to use the residuals from the regression of product coupon with a fixed effect corresponding to each set of dominated products.

FIGURE 4: DISPERSION IN REVERSE CONVERTIBLE RETURNS (DOMINATED/SUPERIOR)

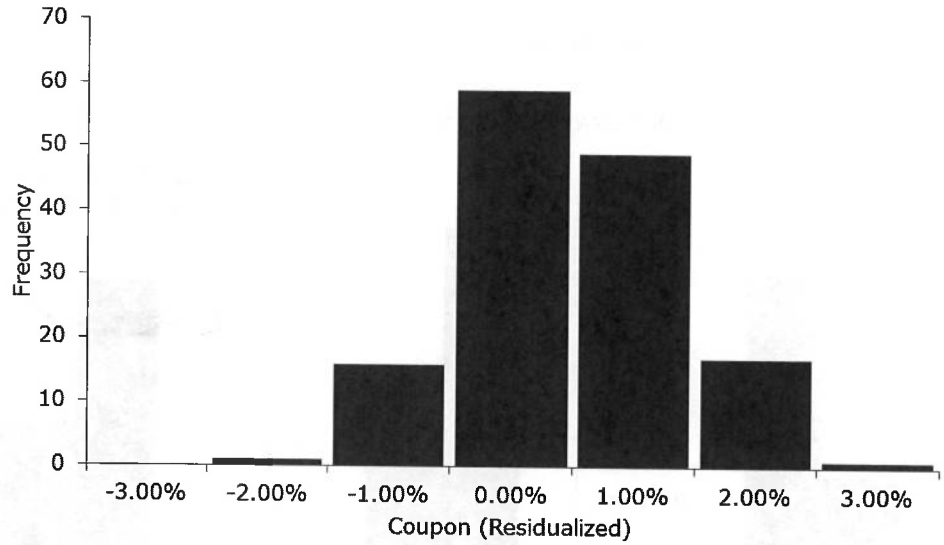


Figure 4 Notes: The figure displays dispersion in reverse convertible coupons among those convertibles that are either dominated or superior in the data set. Figure 4 plots coupon dispersion within dominated/superior products. I define a set of dominated/superior products as all products with the same issue date, issuer, price, and convertible payout. Figure 4 plots the distribution of $c_{i,j} - \bar{c}_j$. Here, i indexes the reverse convertible issuance and j indexes a set of dominated/superior products.

Figure 4 suggests a substantial amount of "real" coupon dispersion exists among reverse convertibles. The dispersion in Figure 4 cannot be a function of some unobserved product characteristic, because these products are otherwise identical. The standard deviation of returns within dominated/superior products is 0.90%. This result suggests unobserved product characteristics do not drive the bulk of the dispersion observed earlier in Figure 3.

4.2 Demand for Dominated Products

The market for reverse convertibles exhibits a substantial amount of "real" dispersion in returns/coupons that unobserved product heterogeneity cannot explain. The economic importance of the dispersion hinges on how often consumers purchase the inferior reverse convertibles.

Figure 5 Panels A-C plot the average characteristics of the dominated and superior reverse convertibles discussed in the previous section. On average, the coupon and subsequent return of the superior reverse convertible is 1.60% points higher than the corresponding dominated reverse convertible. Panel B indicates that, on average, consumers collectively purchased 16% more of the dominated product. Not only are consumers buying dominated products, but they are also actually purchasing more of them relative to the superior product.

FIGURE 5: DOMINATED AND SUPERIOR PRODUCTS

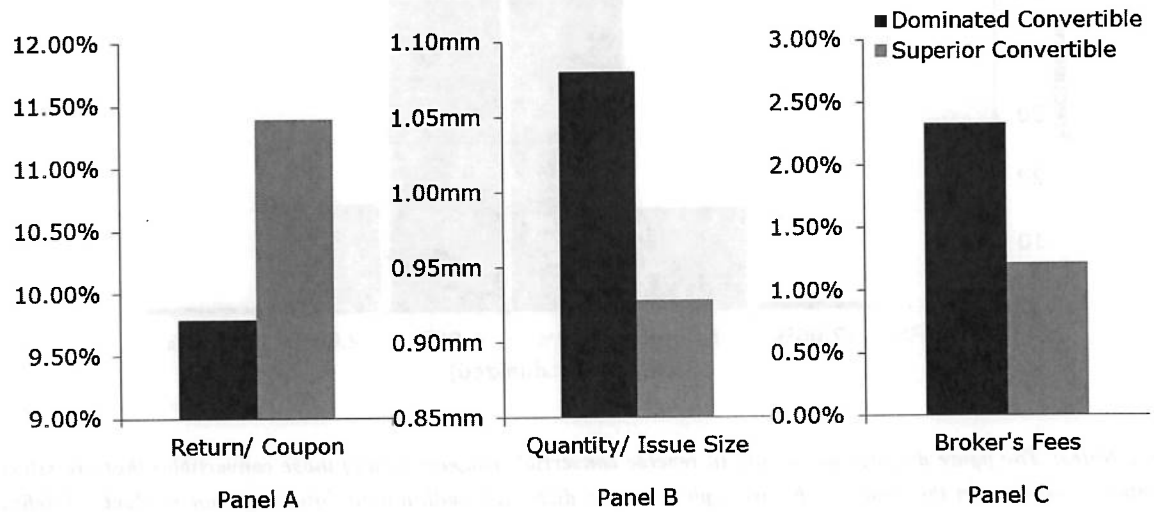


Figure 5 Notes: The figure displays the average characteristics of all of the dominated and superior reverse convertibles in the data set. The data set covers all US-issued, SEC-registered, one-year reverse convertible bonds. A reverse convertible is defined as dominated if a reverse convertible exists with the same issuer, issue date, price, underlying equity, and principal payment structure, with a higher fixed-rate coupon.

The result that consumers purchase more of the dominated product is critical because a standard search model would not predict this finding. In fact, a standard search model would predict the exact opposite: consumers should purchase more of the superior product. Consider a simple example in which two products exist, with one clearly superior to the other. In a simple undirected search model, consumers find each product with equal probability. When a consumer searches and finds the superior product, he simply purchases it and stops searching. If a consumer searches and finds the inferior product, consumers with low search costs continue searching for the superior product, whereas consumers with high search costs purchase the inferior product. Hence, consumers will purchase more of the superior product, provided consumers see both products with equal probability.[17]

Figure 5 illustrates the main points of the empirical and theoretical analysis of the paper. Panel A suggests the consumer's investment problem is fundamentally a search problem. I argue

[17]See Hortaçsu and Syverson (2004) for another example with a simple directed search model.

that consumers buy inferior reverse convertibles simply because they are not aware of the better convertible. However, Panel B suggests there may be more to the story beyond a simple search model. Consumers collectively purchase more of the dominated product. Last, Panel C shows the average fee paid to brokers for selling reverse convertibles. On average brokers, earned a 1.12% point higher commission for selling the dominated product. I argue and show more formally in the preceding section that the brokers and the incentives of brokers play a critical role in determining demand for reverse convertibles.

4.3 Product Distribution, Fees and Demand for Reverse Convertibles

The results from sections 4.1 and 4.2 indicate consumers often fail to purchase the best available product. I argue consumer search in conjunction with broker intermediation rationalizes such behavior. Brokers earn large and heterogeneous fees for selling different reverse convertibles. Figure 6 displays the distribution of broker's fees. Product issuers incentivize brokers with high fees to sell certain products.

FIGURE 6: BROKER'S FEES

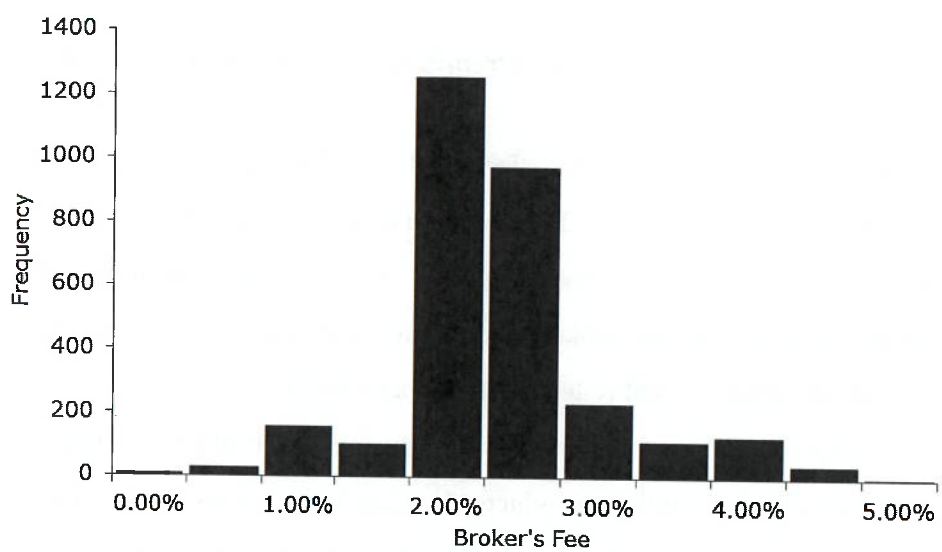


Figure 6 Notes: The figure displays the distribution of fees paid by issuers to brokers.

In this section, I analyze the relationship between product fees and demand for reverse convertibles. First, I estimate several reduced-form specifications to characterize the relationship between product issuance, broker's fees, and other product characteristics. I then look at the relationship between broker's fees and other product characteristics. In general, the incentives of brokers do not appear to align with consumers. All else equal, consumers purchase more of high-fee product. Furthermore, products with high fees have worse payoffs on average.

4.3.1 Issuance Size vs. Product Characteristics

I first analyze the relationship between reverse convertible bond issuances and product characteristics. As will be discussed in detail, under certain assumptions, the following specifications and analysis can be interpreted in a simple linear demand framework.

Theoretically, both consumers and product issuers should value reverse convertibles based solely on their risk and return. In other words, under a risk-neutral framework, a reverse convertible should be valued based on its coupon, issuer credit risk, and embedded equity put option. I rely on two specifications to examine the relationship between product characteristics and issuance. I first regress bond-issuance size on the product-specific coupon, fee, embedded option premium, and issuer CDS spread:

$$Size_j = \beta Fee_j + \alpha Coupon_j + \gamma^{Opt} Option\_Premium_j + \gamma^{CDS} CDS_j + Fixed\_Effects \qquad (2)$$

I also include issuer, month, and equity fixed effects and control for the type of reverse convertible issued (single vs. continuous observation). The observations are reverse convertible bond issuances such that j indexes a particular reverse convertible bond. One of the key variables of interest is the relationship between issuance size and broker's fee. Recall that conditional on the risk and return of a product, consumers should be ambivalent over the broker's fee.

As a robustness check, I estimate a corresponding demand specification in which I restrict the data set to the set of dominated/superior products. I estimate the regression of quantity issued on broker's fees and coupon. I also include a fixed effect for each set of dominated/superior reverse convertibles:

$$Size_j = \beta Fee_j + \alpha Coupon_j + Fixed\_Effects \qquad (3)$$

The fixed effect captures all other product characteristics other than the product fee and coupon.

The linear specifications (1) and (2) help summarize the data. Under certain restrictions/assumptions, one could choose to interpret the above specifications as linear demand estimates. The main concern in interpreting either equation (1) or (2) in a causal demand framework is the potential endogeneity of the right-hand-side variables. One advantage of studying reverse convertibles is that in a risk-neutral framework, the fee, coupon, CDS, and option premium should capture all the relevant characteristics of a reverse convertible.[18] Furthermore, when I restrict products to the set of dominated/superior products (eq. 3), I am able to control for all product characteristics. Also, the issuer typically sets the product characteristics one month in advance of a sale. For these reasons, any unobserved error term is likely to be idiosyncratic and uncorrelated with product characteristics. In some sense, the reverse convertible market offers a near perfect setting to study demand because I see identical products being sold at the same time at different parts of the demand curve.

Table 2 displays the regression estimates corresponding to equations (2) and (3). Columns (1)-(4) include the results for the full data set, whereas columns (5) and (6) display the regression results corresponding to when the data set is restricted to the dominated/superior products. The relevant coefficients not only have the expected sign, but are also statistically significantly different from zero. As expected, the product-issue size is positively correlated with coupon and negatively correlated with equity option premium and issuer credit risk. The results from column (3) indicate that a one percentage point increase in coupon is associated with a \$110,800 increase in issue size. Similarly, a one percentage point increase in equity put option premium is correlated with a \$61,800 decrease in issue size, whereas a one percentage point increase in issuer credit risk (CDS spread) is correlated with a \$431,400 decrease in issue size. Under a risk-neutral framework, one would expect consumers to trade off coupon one for one with both CDS spread and option premium, such that $\alpha = -\gamma^{Opt} = -\gamma^{CDS}$. Overall, the results suggest consumers trade off coupon and option premium roughly one for one but appear relatively averse to credit risk. One potential explanation for this finding is that the data set covers the peak and aftermath of the 2008 financial crisis. With the collapse of Lehman Brothers and Bear Sterns, consumers may have been more sensitive to the default risk of investment banks.

[18]One might also be concerned with the potential marketing/advertising of these products. SEC regulations require that any special marketing materials be filed with the SEC. Given the cost of marketing these products and the relatively short offering periods, no special marketing materials were filed for any of the reverse convertibles in the sample. Each convertible was marketed to consumers with the prospectus, which formally lays out the details of the security.

TABLE 2: ISSUE SIZE VS. PRODUCT CHARACTERISTICS

Dependent Variable	Size	ln(Size)	Size	ln(Size)	Size	ln(Size)
Broker's Fee	44.72***	28.48***	43.34***	28.95***	14.23*	50.02***
	(8.26)	(4.70)	(8.84)	(5.01)	(7.91)	(17.37)
Coupon	12.88***	12.57***	11.08***	10.20***	-4.31	6.10
	(3.08)	(1.61)	(3.49)	(1.77)	(5.55)	(15.21)
Option Premium	-7.64***	-7.51***	-6.18**	-5.69***		
	(2.17)	(1.25)	(2.49)	(1.35)		
CDS Spread			-43.14*	-47.57***		
			(23.39)	(14.34)		
Continuous Obs.	-3.66***	-2.16***	-4.01***	-2.08***		
	(0.34)	(0.15)	(0.45)	(0.18)		
Dominated Products					X	X
Observations	3,066	3,066	2,680	2,680	143	143
R-squared	0.484	0.657	0.475	0.667	0.716	0.726

*Table 2 Notes: Table 2 displays the results from the regressions of quantity issued and broker's fees on the specified variables (eq. 2 and 3). Each specification includes issuer, underlying equity, and month fixed effects. Continuous observation is an indicator variable indicating the reverse convertible is a continuous rather than a single observation reverse convertible. Coupons and fees are measured such that 0.10 corresponds to a 10% coupon/fee. Quantity issued is measured in millions. Huber-White robust standard errors are reported in parentheses. *,**, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.*

The regression results indicate demand is increasing in coupon and decreasing in CDS spread and option premium, but also that demand is increasing in broker's fees. In each specification, I estimate a positive and significant relationship between broker's fees and issue size, even when I restrict the data set to the set of superior/dominated products. The results from column (1) indicate that a one percentage point increase in broker's fees is correlated with a $447,200 increase in issue size. Recall that conditional on the risk and return of a product, consumers should be apathetic toward broker's fees. One might be concerned that some omitted product characteristic that is positively correlated with fees and size might drive this relationship. However, I am able to control for all product characteristics, especially when I restrict the data set to the set of dominated/superior products. These results suggest that brokers are more inclined to sell high-fee products.

4.3.2 Fees vs. Product Characteristics

The estimation results from the Table 1 suggest that all else equal, consumers buy more products with higher broker's fees. Because consumers are theoretically unaffected by the broker's fee, these results suggest brokers are directing consumers to higher-fee products. This finding raises concerns over the conflict of interest between brokers and consumers, especially if products with higher fees have lower returns and higher option premiums and issuer default risk. I examine this relationship further by estimating the following specification in which I regress the broker's fee on the set of product characteristics:

$$Fee_j = \beta_1 Coupon_j + \beta_2 Option\_Premium_j + \beta_3 CDS_j + Fixed\_Effects \tag{4}$$

I also include issuer, month, and equity fixed effects and control for the type of reverse convertible issued. Estimated coefficients $\beta_1 < 0$ and/or $\beta_2 > 0$, $\beta_3 > 0$ would be indicative of a conflict-of-interest problem.

As a robustness check, I again restrict the data set to dominated/superior products and regress broker's fees on the product coupon, and include a fixed effect for each set of dominated/superior reverse convertibles:

$$Fee_j = \beta_1 Coupon_j + Fixed\_Effects \tag{5}$$

Restricting the data set again to dominated/superior products should help limit any concerns over the endogeneity of coupon and product fees.

Table 3 displays the estimation results corresponding to equations (4) and (5). The columns differ in terms of which co-variates are controlled for, whether the regression results are weighted by the square root of the issuance size, and the data set used. The results indicate that fees are negatively correlated with product coupon and positively correlated with equity option premium and issuer credit risk. The estimated coupon coefficients in all six specifications are negative and significant at the 1% level. The estimates indicate a one percentage point increase in coupon is associated with a 0.11% decrease in product fees. Similarly, a one percentage point increase in option premium is correlated with a 0.07% increase in product fees. The estimates from column (4) imply a one percentage point increase in the issuer's CDS spread (issuer credit risk) is correlated with a 0.08% increase in product fees. Although the magnitude of the estimated coefficients is relatively small, the average level of fees in the data set is 2.20%.

TABLE 3: BROKER'S FEES VS PRODUCT CHARACTERISTICS

Variables	(1)	(2)	(3)	(4)	(5)	(6)
Coupon	-0.10***	-0.11***	-0.10***	-0.10***	-0.53***	-0.68***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.09)	(0.07)
Option Premium	0.07***	0.08***	0.07***	0.07***		
	(0.01)	(0.01)	(0.01)	(0.01)		
CDS Spread			0.06	0.08*		
			(0.04)	(0.05)		
Continuous Obs.	0.01***	0.00***	0.01***	0.00***		
	(0.00)	(0.00)	(0.00)	(0.00)		
Weighted		X		X		X
Dominated Products					X	X
Observations	3,066	3,066	2,680	2,680	143	143
R-squared	0.600	0.613	0.614	0.620	0.707	0.833

*Table 3 Notes: Table 3 displays the results from the regressions of broker's fees on the specified variables (eq. 4 and 5). Each specification includes issuer, underlying equity, and month fixed effects. The weighted specifications are weighted by the square root of the issuance size. Continuous observation is an indicator variable indicating the reverse convertible is a continuous rather than a single observation reverse convertible. Coupons and fees are measured such that 0.10 corresponds to a 10% coupon/fee. Huber-White robust standard errors are reported in parentheses. *,**, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.*

Overall, the results from the empirical analysis confirm the existing concerns in the literature that the incentives of brokers do not align with the incentives of consumers. All else equal, consumers are more likely to buy reverse convertibles with high broker's fees. However, reverse convertibles with high fees tend to have worse payoffs. In this sense, brokers are incentivized to sell consumers inferior products.

5 Model

The heterogeneity in reverse convertible risk-adjusted returns raises the question: why do consumers buy inferior convertibles and why do issuers and brokers create and sell both good and bad convertibles? Furthermore, why are consumers actually more likely to purchase dominated products? This section develops a dynamic discrete time model of financial distribution that rationalizes consumer and broker behavior. The model is then structurally estimated and used to analyze and quantify the economic implications of the proposed broker regulations.

The key features of the model are motivated by the preceding empirical analysis and features of the reverse convertible market. The prevalence of dominated financial products suggests that the consumer's investment problem is fundamentally a search problem. Consumers buy dominated products simply because they are unaware of or unable to purchase better alternatives. In the model, consumers sequentially search over the product space with the aid of a broker. Brokers select a product to show each client based on the corresponding product specific broker's fee weighted by the probability the client purchases the product. In selecting products for a client, the objective of a broker is to maximize brokerage commissions rather than to maximize consumer utility. This formulation is supported by the results from the previous empirical section. Lastly, brokers utilize the product space to price discriminate across consumers, showing high fee dominated products to unsophisticated consumers and low fee superior products to sophisticated consumers. The key innovation in the model is that brokers endogenously direct the search of consumers according to the incentives of brokers and a broker's ability to price discriminate across consumers.

5.1 Model Overview

The model involves three types of market participants: consumers, brokers (serving as financial intermediaries) and product issuers. Although largely applicable to most retail financial products, the model is tailored to the distribution of reverse convertible bonds. Product issuers create reverse convertibles and then sell them through brokers to consumers. The actions of issuers and set of available reverse convertibles are taken as given. Rather, the model focuses on the endogenous interactions between brokers and consumers, taking the product space as given. Reverse convertibles are characterized by their payoff c (coupon), short equity put premium e, product issuer credit risk d, and broker's fee/commission f. All bonds are all sold at a fixed par price of 100%. Thus the quadruplet (c, e, d, f) defines a financial product.

Each consumer possesses demand for exactly one reverse convertible bond. Consumers sequentially search over the product space one product at a time. Brokers direct the search process of consumers, informing consumers of the available products. Each period, a broker chooses which reverse convertible bond to show her consumer client. Brokers only show one reverse convertible to each consumer at a time. The consumer elects to either purchase the bond offered or continue searching for a new investment opportunity next period. Consumers can only purchase products offered to them by brokers. If the consumer purchases the bond j, he receives utility flow $\mathcal{U}(c_j, e_j, d_j)$ and his broker receives a fee f_j, that is paid by the product issuer. If the consumer decides to continue searching, he is matched with a new broker and is offered a new product next period.

5.2 Consumer Behavior

Each consumer must purchase exactly one reverse convertible bond. Consumers value financial products based on their risk and return. Product j with return c_j (coupon), put premium e_j, and issuer credit risk d_j generates consumer utility $u_j = \mathcal{U}(c_j, e_j, d_j)$. Utility is increasing in return and decreasing in put premium and issuer credit risk such that $\mathcal{U}_c > 0$, $\mathcal{U}_e < 0$ and $\mathcal{U}_d < 0$. The utility function is specified as a linear function of return/coupon, equity put premium, and issuer credit risk. Issuer credit risk is measured using the corresponding one year CDS spread.

$$u_j = \alpha Coupon_j + \gamma^{Opt} Option\_Premium_j + \gamma^{CDS} CDS_j \tag{6}$$

This utility formulation is roughly consistent with the risk neutral fair value of a reverse convertible. If consumers value reverse convertibles according to the risk neutral prices, consumers should be willing to trade off coupon and equity put premium and issuer credit risk roughly one for one such that $\alpha = -\gamma^{Opt} = -\gamma^{CDS}$ (assuming no discounting).

There are two important things to note regarding the utility formulation. First, neither the price of a reverse convertible nor the broker's fee enters the consumer's utility function. This is because all reverse convertibles are sold at a fixed price of par (100%). The broker's fee is paid by the product issuer rather than the consumer. In this sense, the broker's fee represents the portion of profits shared between the issuer and the broker. Conditional on the risk and return of a product, consumers are apathetic regarding the broker's fee.

Second, the utility formulation implies that the products are vertically rather than horizontally differentiated. Notice that the utility specification does not include an unobserved product and consumer specific error term. Consequently, consumers possess a clear rank ordering over the product space.

Costly search prevents consumers from always simply purchasing the product yielding the highest utility. Each period, consumers must pay a search cost v_i in order to observe a new product offer from a broker. Search costs reflect the time and effort of interacting with a broker as well as the forgone opportunity cost of delaying investment. Search costs for a given consumer are constant over time but are heterogeneous across consumers. Consumers are classified into two types based on their search costs: low search cost/sophisticated investors and high search cost/unsophisticated investors.[19] The fraction of high search cost/unsophisticated investors in the population is denoted ω_H. Search costs among sophisticated/low cost consumers are distributed $v_i \sim F_L(\cdot)$ while search costs among unsophisticated/high cost consumers are distributed $v_i \sim F_H(\cdot)$. Without any loss in generality, consumer types are defined such that the average search cost of sophisticated investors is lower than the average search cost of unsophisticated investors $\mathrm{E}_L[v] < \mathrm{E}_H[v]$. Consumer types reflect the information observed by brokers. Brokers observe a consumer's type and preferences but not his exact search cost. Thus brokers have incomplete information regarding the exact level of financial sophistication of each customer. Neither search costs nor consumer type are observed by the econometrician. As shown in the proceeding section, brokers will select different products to show different consumer types and will essentially price discriminate across sophisticated and unsophisticated investors.

Consumers sequentially search for the optimal investment among the discrete product space $\{u_1, u_2, ...u_n\}$. Products are numbered such that $u_j \leq u_{j+1}, \forall j$. While searching, a consumer receives an offer from a broker each period and then must elect to either purchase the offered bond or continue searching. If the consumer decides to continue to search he pays a search cost v_i and receives an offer from a new broker in the preceding period. All subsequent product offers are drawn i.i.d. from either the stationary distribution $H_L(\cdot)$ or the stationary distribution $H_H(\cdot)$ depending

[19]Both types of consumers are assumed to be completely rational and differ with respect to search costs. This is opposed to previous research, such as Gabaix and Laibson (2006), that allow for consumer heterogeneity where a subset of consumers are naive. The heterogeneity in search costs relates to the growing literature on financial literacy such as Hastings and Tejeda-Ashton (2008) Choi et al. (2010).

on whether the consumer is a sophisticated or unsophisticated investor. As will be discussed in the proceeding section, the key innovation in the model is that the distribution of products observed by low type and high type consumers $H_L(\cdot)$ and $H_H(\cdot)$ are endogenously determined based on the incentives of brokers. In equilibrium, consumer beliefs about $H_L(\cdot)$ and $H_H(\cdot)$ are correct and completely rational. As discussed in the proceeding section, I focus on stationary equilibrium in which the distribution of offered products is constant over time.[20] I abstract away from the broker/consumer matching process by assuming that conditional on type, brokers and consumers are ex-ante identical and are randomly assigned.

Let $V(u_j, v_i, T)$ denote the value function of a consumer with search cost v_i of type T (either low, L, or high, H) that is offered a product yielding utility u_j. A consumer offered product j can either purchase the product or pay a search cost and continue searching. Formally the consumer's problem is[21]

$$V(u_j, v_i, T) = \max \left\{ u_j, -v_i + \sum_{k=1}^{n} \rho_{k,T} V(u_k, v_i, T) \right\}$$

Purchasing the product j yields utility flow u_j while the expected utility of searching is $-v_i + \sum_{k=1}^{n} \rho_{k,T} V(u_k, v_i, T)$. Here $\rho_{j,T}$ reflects the probability a consumer of type T observed product j. The sets of offering probabilities $\rho_{1,L}, \rho_{2,L}, ..., \rho_{n,L}$ and $\rho_{1,H}, \rho_{2,H}, ..., \rho_{n,H}$ are endogenously determined based on incentives of brokers. Collectively the probabilities $\rho_{1,L}, \rho_{2,L}, ..., \rho_{n,L}$ and $\rho_{1,H}, \rho_{2,H}, ..., \rho_{n,H}$ form the distributions $H_L(\cdot)$ and $H_H(\cdot)$.

Under this framework, consumers optimally search by adopting a reservation utility.[22] A consumer with search cost v_i of type T optimally searches until he is shown an investment product that exceeds his reservation utility $u_T^r(v_i)$. Consumers will optimally continue searching as long as the consumer's expected benefit of search is greater than his search cost. Suppose a consumer is offered product yielding utility u_j, the consumers expected benefit of search is given by $\sum_{k=j+1}^{n} \rho_{k,T}(u_k - u_j)$ which is equal to the probability the consumer sees a better product than u_j weighted by the gain

[20] Alternatively one can think of the market as clearing instantaneously.

[21] The equivalent formulation with a continuous product space is given by

$$V(u_j, v_i, T) = \max \left\{ u_j, -v_i + \int_{\underline{u}}^{\bar{u}} V(u', v_i, T) dH_t(u') \right\}$$

where $[\underline{u}, \bar{u}]$ is the support of available products.

[22] See McCall 1970, Rogerson et al. 2005, and Hortaçsu and Syverson (2004) for a further discussion of search problems and a proof of the optimal strategy.

in terms of utils.[23] The optimal strategy is then[24]

$$\text{Continue Searching:} \quad \underbrace{v_i}_{\text{Cost}} \leq \underbrace{\sum_{k=j+1}^{n} \rho_{k,T}(u_k - u_j)}_{\text{Expected Benefit}}$$

$$\text{Purchase:} \quad v_i \geq \sum_{k=j+1}^{n} \rho_{k,T}(u_k - u_j) \tag{7}$$

The reservation utility is equal to the utility generated by product j, $u_T^r(v_i) = u_j$, such that $\sum_{k=j+1}^{n} \rho_{k,T}(u_k - u_j) \leq v_i \leq \sum_{k=j}^{n} \rho_{k,T}(u_k - u_{j-1})$. A consumer purchases the product if it exceeds his reservation utility, $u_T^r(v_i)$, otherwise he continues searching. An individual's optimal reservation utility, $u_T^r(v_i)$ is a weakly decreasing function of his search cost v_i. A consumer with zero search costs searches until he finds the product yielding the highest utility u_n while a consumer with infinite search costs simply selects the first product offered. Since brokers observe a consumer's type, the distribution of product offered $H_L(\cdot)$ and $H_H(\cdot)$ will likely vary across types in equilibrium. Consequently, the optimal reservation strategy of an individual with search cost v_i will also depend on the agent's type T. The analysis focuses on stationary equilibrium such that the distribution of consumer types and search costs among agent types is at a steady state in the searching population. Let $G_L(\cdot)$ and $G_H(\cdot)$ denote the stationary distribution of reservation utilities among low and high cost consumer types in equilibrium.

A couple of underlying assumptions in the model are worth noting. In the model framework, consumers know the distribution of product offerings $H_L(\cdot)$ and $H_H(\cdot)$ (or equivalently $\rho_{j,T}\ \forall j, T$) but are unable to purchase a product without the aid of the broker. Although not applicable to all financial markets, this framework seems reasonable in the setting of reverse convertible bonds. Reverse convertible bonds have short marketing periods (typically less than one month) and are SEC registered products which makes them costly to market directly to end consumers. Consequently, issuers do not market these products directly to consumers. The prevalence of dominated

[23] This formulation assumes consumers can recall and purchase products observed in prior periods; however, in practice consumers will never find it optimal to do so.

[24] The equivalent optimal reservation strategy in the formulation with a continuous product space is given by

$$v_i = \int_{u_T^r}^{\bar{u}} (u' - u_T^r) dH_T(u')$$

where $[\underline{u}, \bar{u}]$ is the support of available products.

products indicates that search is a key component of the consumer's problem. Investor suitability regulations (FINRA Rule 2111) require that reverse convertible investors meet a certain level of financial sophistication, risk tolerance etc.. Hence, even though reverse convertible investors may not know the exact distribution of product offerings, they may still have realistic expectations over the distribution of product offerings based on previous experience and the prices of more transparent assets.

5.3 Broker Behavior

Brokers act as a liaison between the end consumers and the financial product issuers. Brokers observe the full scope of available products. Each period brokers offer each consumer an individual specific financial product tailored to the consumers level of sophistication/type. If the consumer purchases the product, the product issuer pays the corresponding broker a product specific fee. Fees f_j for a given product j are fixed but are heterogeneous across products.

Each issuer creates a suite of financial products available to and observed by all of the brokers. Let $\mathcal{J} = \{u_1, u_2, ...u_n\}$ denote the product space available to brokers. For each of her clients, the broker selects the product that maximizes her expected profits

$$\max_{j \in \mathcal{J}} E[\pi_{i,j}] \tag{8}$$

Offering product j to client i, yields an expected profit equal to the probability client i purchases product j multiplied by the returns from selling product minus the cost of offering the product. Brokers observe the preferences and types of their clients but do not observe each client's specific search cost. Recall that a consumer purchases a product if it exceeds his reservation utility. The probability product j exceeds a type T consumer's reservation utility and thus the probability a consumer purchases the product is given by $G_T(u_j)$. The distribution of reservation utilities $G_T(\cdot)$ are endogenously determined based on the consumers optimal search strategy (12). The expected profit of offering product j to client i of type T is then

$$\mathrm{E}[\pi_{i,j,T}] = f_j G_T(u_j) + \eta_{i,j} \tag{9}$$

where $f_j G_T(u_j)$ is the broker's expected revenue and $\eta_{i,j}$ is a product/consumer specific marketing cost incurred by the broker. The cost term $\eta_{i,j}$ is unobserved (by the econometrician) and is assumed to be distributed T1EV. The expected profit of showing product j is increasing in the fees associated with the product and the utility generated by the product. The better the product, the higher the probability it will exceed a consumers reservation utility.

A key assumption in the model framework is that brokers only show a client one product at a time and that each particular broker and client interact at most one time. These assumptions rule out any learning between brokers and clients. For tractability reasons, these assumptions simplify the broker's profit maximization problem to a static problem while the consumer's search problem remains dynamic. In practice these assumptions may be reasonable when applied to the reverse convertible setting. It seems unlikely that a broker would simultaneously show a superior and dominated product to a client. Simalarly, a broker may be hesitant to show a client a superior product in a proceeding period after first showing them a dominated product or vice versa.

The probability that a broker selects product j to offer to a client of type T, denoted $\rho_{j,T}$, is given by

$$\rho_{j,T} = \Pr\left(\mathrm{E}[\pi_{i,j,T}] > \mathrm{E}[\pi_{i,k,T}] | \forall k \in \mathcal{J}_{-j}\right)$$

Given the distributional assumption of the cost shock $\eta_{i,j}$, the probability that a broker selects product j follows the multinomial logit distribution

$$\rho_{j,T} = \frac{exp\left(f_j G_T(u_j)\right)}{\sum_{k=1}^{n} exp\left(f_k G_T(u_k)\right)} \tag{10}$$

The offering probabilities $\rho_{j,T}$ for the various set of products generate the distribution of available products $H_L(\cdot)$ and $H_H(\cdot)$ observed by the two types of consumers. Note that distribution of reservation utilities $G_L(\cdot)$ and $G_H(\cdot)$ and the distribution of product offerings $H_L(\cdot)$ and $H_H(\cdot)$ are endogenously and simultaneously determined in equilibrium according to optimal consumer and broker behavior described in equations (7) and (8).

The probability that a broker shows a particular product to a client is a function of the product's fees as well as the probability that the client purchases the product. All else equal, the probability that a broker selects a particular product to show a client is increasing in the product fees

$$\frac{\partial \rho_{j,T}}{\partial f_j} = G_T(u_j)\rho_{j,T}(1-\rho_{j,T}) > 0$$

Brokers only earn the fee if the consumer purchases the product. The better the product offered, the more likely consumers are to purchase the product. For this reason, the probability a broker selects a particular product to show a client, all else equal, is increasing in the utility generated by the product

$$\frac{\partial \rho_{j,T}}{\partial u_j} = f_j g_T(u_j)\rho_{j,T}(1-\rho_{j,T}) > 0$$

where $g_T(\cdot)$ is the density corresponding to the distribution $G_T(\cdot)$. In this sense, the incentives of brokers and consumers are not totally misaligned. If the fees, f, and costs, η, were fixed across products, brokers would be incentivized to always offer products that generate the highest utility. However, the reduced form results from Section 4.3.2 suggest that fees and product utility are negatively correlated. Overall, consumers are more likely to observe products with higher fees and that generate higher utility.

5.4 Equilibrium

I study a stationary pure strategy Bayes Nash equilibrium. In equilibrium consumers optimally search by employing the reservation strategy described by equation (7). Furthermore, consumer beliefs over the distribution of indirect utilities offered for high and low types, $H_L(\cdot)$ and $H_H(\cdot)$, reflect the true distribution of product offerings generated from broker profit maximization. In equilibrium brokers maximize profits according to equations (8) and (9) where their beliefs over the distribution of reservation utilities reflect the true distributions generated by equation (7), $G_L(\cdot)$ and $G_H(\cdot)$. The distribution of products observed by consumers, $H_L(\cdot)$ and $H_H(\cdot)$, and the distribution of reservation utilities, $G_L(\cdot)$ and $G_H(\cdot)$, are endogenously and simultaneously determined in equilibrium.

The distribution of search costs and consumer types in the population, market parameters and characteristics of available products are all assumed to be constant over time. Or alternatively, the market is assumed to clear instantaneously. The equilibrium is therefore stationary. Consequently, the distribution of product offerings and reservation utilities are constant over time.

6 Model Estimation

The search model described in Section 5 lends itself to structural estimation. Using the reverse convertible data set, I structurally estimate the search model. The model and estimation procedure most closely resembles that of Hortaçsu and Syverson (2004) and Hong and Shum (2006). The key parameters of interest are consumer preferences, the broker's profit functions, and the distribution of reservation utilities, consumer types and search costs.

The model is estimated using the reverse convertible market share level data described in Section 3. Each month and underlying equity defines a reverse convertible market and corresponding market share. For example, all one year reverse convertibles linked to Apple Inc. issued in December 2012 constitute a market. In total there are 423 markets with 1227 different reverse convertibles.[25]

The model is estimated via maximum likelihood. The probability a consumer purchases product j is equal to the probability the broker shows the product to a consumer multiplied by the probability that the product's utility exceeds the consumer's reservation utility. The probability a consumer observes and purchases a bond depends on his consumer type which is observed by brokers but not the econometrician. Thus the probability that consumer i purchases product j is given by

$$\begin{aligned}
\Pr(D_{ij}=1) &= \omega_H \rho_{j,H} G_H(u_j) + (1-\omega_H)\rho_{j,L} G_L(u_j) \\
&= \omega_H \frac{exp(\theta f_j G_H(u_j))}{\sum_{k=1}^{n} exp(\theta f_k G_H(u_k))} G_H(u_j) + (1-\omega_H)\frac{exp(\theta f_j G_L(u_j))}{\sum_{k=1}^{n} exp(\theta f_k G_L(u_k))} G_L(u_j)
\end{aligned}$$

where $D_{i,j}$ is a dummy variable indicating that individual i purchased product j. Here the term ω_H reflects the probability a consumer is a high type, the term $\rho_{j,H}$ or $\frac{exp(\theta f_j G_H(u_j))}{\sum_{k=1}^{n} exp(\theta f_k G_H(u_k))}$ reflects the probability that a high type consumer is shown product j, and $G_H(u_j)$ reflects the probability that product j exceeds a high consumer type's reservation utility. I introduce the parameter θ as a scaling parameter. The parameters to be estimated in the model are θ, the utility parameters corresponding to eq. (6), the distribution of reservation utilities $G_H(\cdot)$ and $G_L(\cdot)$ and the distribution of consumer types ω_H. The distribution of consumer types (high and low) are estimated using a discrete mixing distribution similar to Heckman and Singer (1984).

[25]Note that the original sample consists of 3,066 reverse convertibles. Markets consisting of only one reverse convertible are not used in the model estimation procedure.

I estimate the model using market share data. Hence, the dependent variable is the market share for each product which ranges from zero to one.[26] Note that from the market share data, I only observe bond purchases and do not observe individuals who were shown bonds but elected not to purchase them. A common problem related to demand estimation in the industrial organization literature is how to define and quantify the outside good/alternative which in this setting is not purchasing a reverse convertible. I circumvent the outside good issue by simply estimating the observed conditional probabilities. I estimate the model via maximum likelihood where I condition on the probability that a consumer purchased a reverse convertible from that particular market. The corresponding likelihood used to estimate the model is given by

$$\Pr\left(D_{i,j}=1|\sum_{l=1}^{n}D_{i,l}=1\right)=\frac{\omega_H\rho_{j,H}G_H(u_j)+(1-\omega_H)\rho_{j,L}G_L(u_j)}{\sum_{l=1}^{n}\left[\omega_H\rho_{l,H}G_H(u_l)+(1-\omega_H)\rho_{l,H}G_H(u_l)\right]} \tag{11}$$

Estimating the conditional likelihood solves the outside good problem in this setting.

To facilitate estimation, I assume that consumers employ the same set of reservation strategies across all markets conditional on consumer type. In other words, $G_H(\cdot)$ and $G_L(\cdot)$ are assumed to be constant across all markets. This assumption is equivalent to assuming that the distribution of search costs, consumer types and consumer beliefs over $H_H(\cdot)$ and $H_L(\cdot)$ are constant across markets. For example, this implies that consumers searching for Apple linked reverse convertibles and Microsoft linked reverse convertibles hold the same beliefs over the distribution of available reverse convertibles. This assumption provides additional statistical power to estimate $G_L(\cdot)$ and $G_H(\cdot)$, otherwise they would have to be separately estimated for each market. Although this assumption restricts consumer beliefs, it may not be unreasonable to think consumers searching for Apple or Microsoft linked reverse convertibles employ the same strategy. This assumption could also be relaxed in future work.

The model is parametrized as follows. The utility function is specified as a linear function of coupon, option premium and the CDS spread according to equation (6). I also include issuer brand/fixed effects for the five largest issuers: ABN Amro, Barclays, JPMorgan, RBC and UBS. The parameter θ is added as a flexible scaling parameter for the logit error term.

Estimation of the model requires no additional assumptions regarding the parametric form of

[26]As a robustness check shown in the appendix I also re-estimate the model where each observation is weighted by the market size.

$G_L(\cdot)$ or $G_H(\cdot)$. Following Barseghyan et al (2013), I flexibly estimate the distribution functions $G_L(\cdot)$ and $G_H(\cdot)$ using a third order polynomial approximation to $\log G_L(\cdot)$ and $\log G_H(\cdot)$.[27] I estimate polynomial approximations to $\log G_L(\cdot)$ and $\log G_H(\cdot)$ rather than $G_L(\cdot)$ and $G_H(\cdot)$ to ensure that the estimated distributions $\widehat{G_L(\cdot)}$ and $\widehat{G_H(\cdot)}$ are strictly positive. However, I do not restrict the estimated distribution functions to be weakly increasing. The variation in the data helps identify the curvature of the reservation utility functions $G_L(\cdot)$ and $G_H(\cdot)$. However, the scale of $G_H(\cdot)$ and $G_L(\cdot)$ is not identified separately identified from θ in the above likelihood. The scale of $G_L(\cdot)$ and $G_H(\cdot)$ is pinned down by the fact that all consumers purchase the best product which yields utility $\bar{u}$. In other words, no consumer continues searching if observes the best available product. Hence, $G_L(\bar{u}) = G_H(\bar{u}) = 1$.

The underlying data and model separately identifies the consumer utility and broker parameters as well as the observed distribution of reservation utilities. The utility formulation of the model allows for two normalizations. Due to its arbitrary scale and level, I normalize consumer preferences for coupon equal to one and the constant[28] to zero. Under this normalization, the utility parameters can be interpreted in terms of monetary value or percentage return.

Although each parameter of the model is jointly identified through the data, I provide a brief stylized discussion of the intuition behind the identification of the key parameters of the model. The preference parameters γ^{Opts} and γ^{CDS} measure how consumers trade off option premium and issuer credit risk (CDS) relative to coupon. Identification of preferences is best illustrated through the proceeding thought experiment. Suppose we observe a product with fees f, coupon c, and equity option premium e that has market share s. Now suppose we decrease the coupon from to c to c', $c' < c$. The question we are interested in is how much would the option premium have to decrease by from e to e' to keep the market share of the product unchanged at s. The compensating change in option premium identifies how consumers trade off option premium for coupon.

Intuitively, identification of the distribution of reservation utilities $G_L(\cdot)$ and $G_H(\cdot)$ follows closely to that of the preference parameters. The conceptual experiment we would like to be able to

[27]Note that $G_H(\cdot)$ and $G_L(\cdot)$ are estimated using a smooth polynomial function while $G_H(\cdot)$ and $G_L(\cdot)$ are likely non-smooth in practice. Given that the distribution of available products is discrete $H_L(\cdot)$ and $H_H(\cdot)$, then the distribution of reservation utilities $G_H(\cdot)$ and $G_L(\cdot)$ will also be discrete according to the search model described in Section 5.

[28]I included brand fixed effects for the five largest issuers: ABN Amro, Barclays, JPMorgan, RBC and UBS. The constant represents the brand effect for all other issuers.

run is to freely vary the coupon of a product and see how the corresponding product's market share changes, keeping all other products and product characteristics constant. Such variation allows us to trace out the curvature of the distribution of reservation utilities. The scale of $G_L(\cdot)$ and $G_H(\cdot)$ is pinned down by the fact that all consumers purchase the best product which yields utility $\bar{u}$, i.e. $G_L(\bar{u}) = G_H(\bar{u}) = 1$.

The variation in consumer types is identified by variation in the distribution of product offerings across markets. Specifically, the variation in substitution patterns across markets identifies consumer types. Consider a market consisting of one clearly superior bond and one dominated bond in terms of utility. Now suppose an additional inferior bond is introduced into the market. We can identify the proportion of low cost/sophisticated types based on how the market share of the superior bond changes when an additional inferior bond is introduced into the market. If the market share of the superior bond falls dramatically, that suggests those investors who initially purchased the superior reverse convertible were "lucky" unsophisticated/high cost investors. If the market share of the superior bond does not change much, that suggests that those investors who initially purchased the reverse convertible were primarily sophisticated/ low cost investors. Although the preceeding example is a bit stylized, variation in substitution patterns across markets is the key feature of the data that identifies consumer types.

7 Estimation Results and Analysis

7.1 Estimation Results

The maximum likelihood estimates are reported in Table 4. I first estimate the model allowing for only type of consumer and then estimate the model allowing for two types of consumers: high and low cost consumer types. Column (1) displays the estimates for the one consumer type model while column (2) reports the estimates corresponding to the heterogeneous two consumer type model. As expected, the results indicate that consumer utility is decreasing in equity option premium and issuer credit risk (CDS). In both specifications, I estimate a negative and statistically significant relationship utility and the two measures of risk. The results from column (2) indicate that consumers are indifferent between a 1.00% point increase in coupon and a 0.67% point decrease in option premium. Similarly, consumers are willing to trade off a 1.00% point change in coupon for a 5.55% point decrease in the corresponding CDS spread. Recall that under a risk neutral framework consumers should be willing to trade off option premium and CDS spread roughly one-for-one with coupon. Just as with the reduced form results from Section 4, it appears that consumers are particularly sensitive to issuer credit risk.

In the heterogeneous agent model, I also estimate the distribution of consumer types ω_H. I estimate that 98.94% of the population is comprised of high cost/unsophisticated investors while the remaining 1.06% of the population is comprised of low cost/sophisticated investors. The differentiating factor between the two types is how search costs are distributed across types.

TABLE 4: STRUCTURAL ESTIMATION RESULTS

Variables	(1)	(2)
Coupon (α)	1	1
Option Premium (γ^{Delta})	-0.66**	-0.67***
	(0.26)	(0.013)
CDS Spread (γ^{CDS})	-6.59***	-5.55***
	(2.35)	(0.40)
Scaling Parameter (θ)	39.67	45.72
	(24.35)	(29.23)
ω_H		98.94%**
		(0.49%)
Heterogeneous Agents		X
Observations	1,227	1,227
Number of Markets	423	423

*Table 4 Notes: Table 4 displays the maximum likelihood estimation results for the fully specified model. Standard errors are calculated using the observed Fisher Information Matrix. *,**, *** indicate significance at the 10%, 5% and 1% level.*

7.2 Search Costs

The structural model provides additional quantitative insight into the underlying forces driving the market for reverse convertibles. The empirical evidence suggests that costly search prevents consumers from finding the superior reverse convertibles in the market. Using the model estimates, I am able to recover the distribution of search costs which provides us with an opportunity to determine whether or not the estimates are reasonable and/or economically meaningful.

I recover the search cost distributions as follows. First, from the estimation procedure I estimate the distribution of reservation utilities for both types of consumers, $\widehat{G_L}(\cdot)$ and $\widehat{G_H}(\cdot)$. One of the empirical assumptions is that consumers use the same search strategies across markets; hence, $G_L(\cdot)$ and $G_H(\cdot)$ are constant across markets. Consistent with that assumption, I assume that each consumer's belief over the distribution of indirect utilities offered reflect the empirical density of utilities offered, $\widehat{h_H}(\cdot)$ and $\widehat{h_L}(\cdot)$. To calculate $\widehat{h_H}(\cdot)$ and $\widehat{h_L}(\cdot)$, I first calculate the probability that a broker shows each product j to a client. Given the distribution of reservation utilities and corresponding profit parameters, I calculate $\rho_{j,T}$ for each product and consumer type according to equation (10). Given the set of $\rho' s$ for each product and consumer type, I then calculate the density of indirect utilities for observed product offerings for each consumer type $h_L(\cdot)$ and $h_H(\cdot)$ via kernel density estimation giving each observed market equal weight.[29] Lastly, I calculate the distribution of search costs by inverting the equation

$$v_i = \int_{u^r}^{\bar{u}} (u' - u^r) dH_T(u') \tag{12}$$

Note that equation (12) is continuous product space equivalent to optimal reservation utility in the discrete formulation characterized by equation (7). Here I use the continuous product space formulation since consumer beliefs reflect the empirical density of offered products.

Figures 7 and 8 display the estimated distribution of search costs for the single consumer type and two consumer type models. The estimated search costs from the one agent model displayed in Figure 7 suggest that roughly 50% of the population has search costs below 25bps. In other words, it costs the majority of consumers less than 0.25% to receive an offer from a broker. Figure 8 displays the results for the two agent model. The results indicate that virtually all sophisticated investors have essentially zero search costs.[30] Recall that sophisticated consumers make up 1.06%

[29]Here $h_L(\cdot)$ and $h_H(\cdot)$ are the densities corresponding to the distribution functions $H_L(\cdot)$ and $H_H(\cdot)$. I estimate the density of the indirect utility of product offerings for each type of consumer using a Gaussian kernel and giving equal weight to each market. I use select the kernel bandwith according to Silverman's Rule of Thumb.

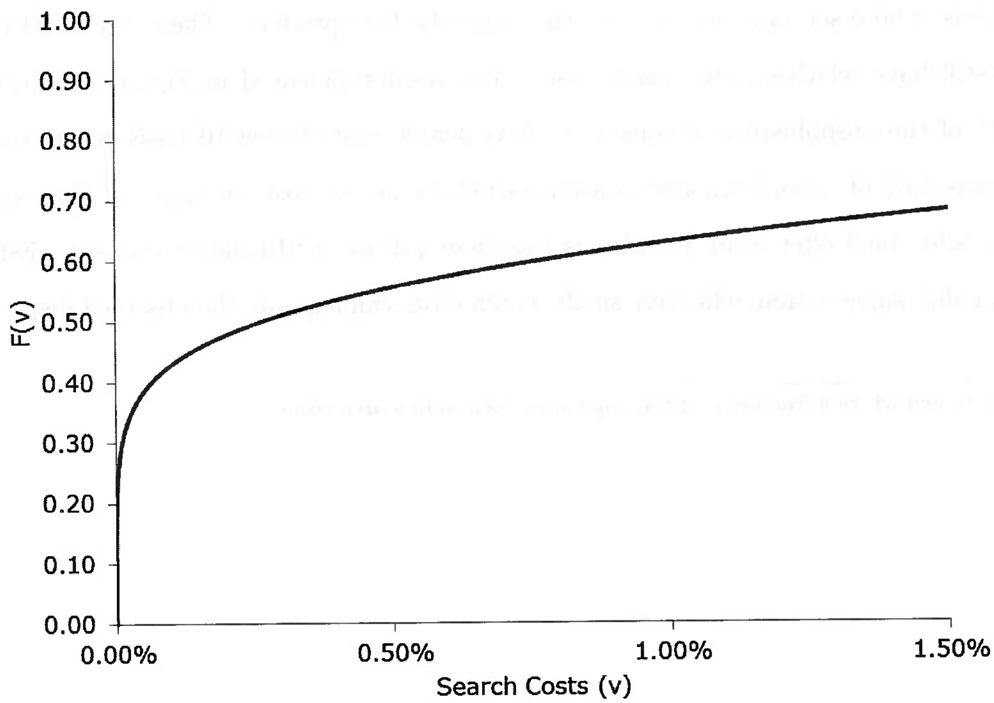
[30]The finding that 1.06% of the population has essentially zero search costs is consistent with Stahl's (1989) model

of the population. This result suggests that roughly one percent of the population is comprised of market mavens, who essentially search until they find the best product. The other 98.94% of the population still have relatively low search costs. The results displayed in Figure 8 indicate that roughly 60% of the unsophisticated consumers have search costs below 10 basis points per offer. In other words 60% of unsophisticated consumers behave as the cost (in terms of time value) of soliciting an additional offer from a broker is less than $10 for a $10,000 investment. Estimates from both models suggest that relatively small search costs can support the observed dispersion in returns.

---

of consumer of search where a fraction μ of the population have zero search costs.

FIGURES 7 AND 8: SEARCH COSTS



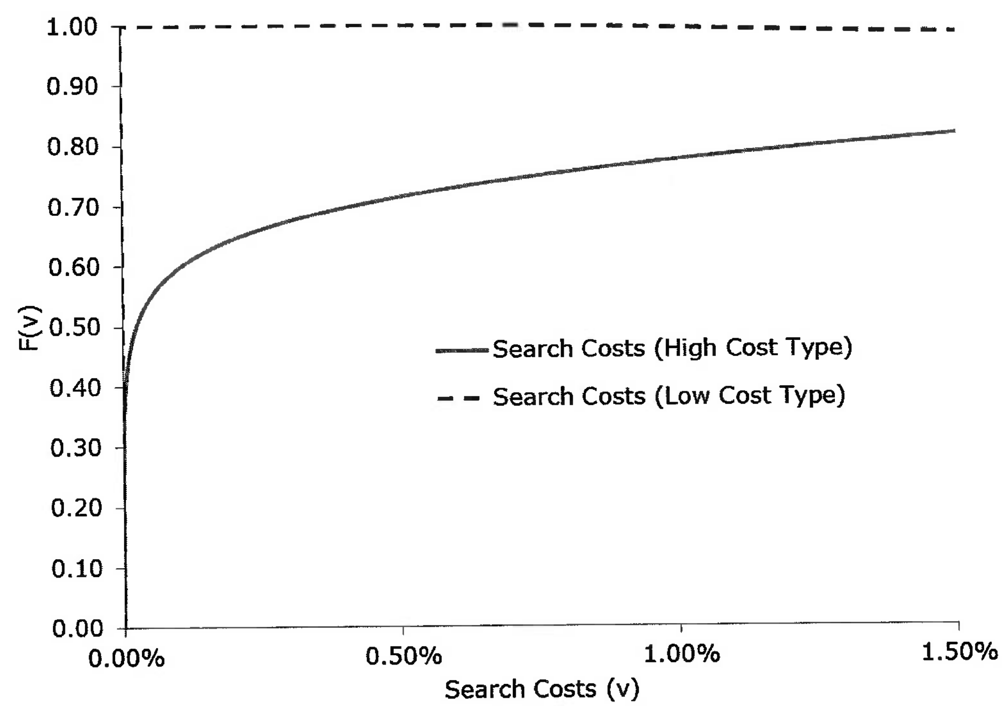


Figure Notes: Figure 7 displays the estimated distribution of search costs corresponding to the single consumer type model. Figure 8 displays the estimated distribution of search costs corresponding to the heterogeneous two consumer type model. The total population is comprised of 98.94% of unsophisticated consumers and 1.06% of sophisticated consumers.

7.3 Broker Behavior

The consumer search problem is compounded by the fact that brokers are not incentivized to show consumers the best available products. The structural estimates help illustrate the incentives of brokers and assess the degree of price discrimination occurring in the reverse convertible market.

Consider the hypothetical market comprised of two nearly identical reverse convertibles where the payout of one dominates the payout of the other. One of the reverse convertibles, the superior reverse convertible, pays a coupon of 12% and a broker's fee of 1.00%. The other reverse convertible, the dominated reverse convertible, pays a coupon of 10% and a broker's fee of 3.00%. We can use the parameter estimates to determine the probability consumers of each type observe each product. Table 5 displays the probability consumers observe each product. Brokers are slightly more likely to show the superior product relative to the dominated product to a sophisticated consumer. However, brokers are almost twice as likely to show an unsophisticated consumer the dominated product relative to the superior product. This helps explain not only why consumers buy dominated products but why consumers are actually purchase more of dominated products.

TABLE 5: IMPLIED SEARCH PROBABILITIES

	Reverse Convertible 1	Reverse Convertible 2
Fee	1.00%	3.00%
Coupon	12.00%	10.00%
Prob. Observed by Sophsiticated Type	0.51	0.49
Prob. Observed by Unsophisticated Type	0.37	0.63

Table 5 Notes: Table 5 displays the probability a broker shows a particular product to a low type and high type consumers in a two product market. Other than the coupon and associated brokers fee, Reverse Convertible 1 and Reverse Convertible 2 are assumed to be identical such that the payout of Reverse Convertible 1 dominates the payout of Reverse Convertible 2. Note that sophisticated consumers are more likely to observe the superior product while unsophisticated consumers are more likely to observe the dominated product.

8 Discussion

Two economic forces/frictions appear to drive the existence and prevalence of dominated products. First, consumers must not be aware of or able to purchase the superior product. I model and argue that the consumer's problem is fundamentally a search problem. Second, the consumers search problem is confounded by the fact that the incentives of brokers do not align with the incentives of consumers. As consumers search for new investment products, they are more likely to see high fee products. Hence the conflict of interest burdens consumers with excess search. The structural estimation results provide a way of quantifying the forces driving consumer behavior in an economically meaningful way. Both economic forces/frictions impact the distribution and total level of consumer and producer surplus. Understanding the costs associated with each force/friction provides insight into optimal financial regulation.

Building on the structural estimation results from the preceding section, I separately analyze the costs and inefficiencies generated by consumer search and the conflict of interest. I first examine the change in total and consumer surplus that would result if we were able to eliminate search costs. I then calculate the expected change in surplus associated if we were able align the incentives of brokers and consumers.

8.1 Search Costs

The fundamental friction in the model is search costs. If search costs were zero, consumers would simply search until they found the best product. Eliminating search costs would remove the market power currently held by brokers and product issuers.

I calculate the change in total search expenditure and the change in expected consumer returns if consumers had zero search costs. As discussed in the preceding section, I can calculate the implied distribution of search costs from the estimated distribution of reservation utilities and implied distribution of product offerings according to equation (12). A consumer's expected search expenditure is equal to his search cost multiplied by his expected number of searches. The expected number of searches follows a geometric distribution and is equal to one divided by the probability a consumer observes a product that exceed his reservation utility. The expected total search expenditure of a consumer with search cost v_i of type T is given by

$$\text{Search Expenditure}(v_i, T) = \frac{v_i}{1 - H\left(u_T^r(v_i)\right)}$$

The average search expenditure among unsophisticated and sophisticated investors is 2.09% and 0.01% respectively. Eliminating such search expenditures represents real surplus gains to the economy.

In the search model framework, the expected return of searching for a product is equal to a consumer's reservation utility $u_T^r(v_i)$. With zero search costs, all consumers would search until they found the best product generating utility $\bar{u}$. Consequently the average expected gain for consumers is given by

$$\Delta Expected\_Return_T = \bar{u} - \int_{-\infty}^{\infty} u' dG_T(u')$$

The term $\int_{-\infty}^{\infty} u' dG_T(u')$ reflects the average reservation utility or simply the average expected return for consumers of type T.

Table 6 reports the average change in search expenditures and average change in risk-adjusted returns if all consumers had zero search costs. On average, the risk-adjusted returns of high cost unsophisticated investors would increase by 5.10% points. Similarly, the risk-adjusted returns of sophisticated investors would increase by 0.01% points. Not surprisingly unsophisticated consumers benefit substantially more from the policy. Just over half of the gain in consumer returns (41%) is attributable to the decline in search expenditures. The remaining gain in consumer returns represents a transfer from brokers and product issuers to consumers.

TABLE 6: ECONOMIC IMPACT OF SEARCH COSTS

	Consumer Type		
	Unsophisticated	Sophisticated	Average Consumer
Avg. Change in Search Expenditure	-2.09%	-0.01%	-2.07%
Avg. Change in Expected Return	5.10%	0.01%	5.05%

Table 6 Notes: Table 6 displays the hypothetical gains to total and consumer surplus if all consumers had zero search costs.

This analysis reflects a partial equilibrium analysis in that the characteristics of available reverse convertibles and the actions of product issuers are fixed. If consumers had zero search costs that would eliminate all market power currently held by brokers and issuers. As issuers adapt to the new zero search cost environment, Nash Bertrand competition among issuers will drive them to create the best possible reverse convertible (in terms of utility) where issuers earn zero markups. In this sense, the estimates from Table 6 provide a lower bound on the change in expected consumer returns.

8.2 Broker Incentives

Brokers select the product that maximizes the broker's expected profit rather than the product that maximizes the utility of consumers. This second economic force burdens consumers with excess search. The preceding structural estimates help determine the cost associated with the asymmetric incentives between brokers and consumers.

Under the preceding framework, the probability a broker shows product j to a client of type T is given by

$$\rho_{j,T} = \frac{exp\left(f_j G_T(u_j)\right)}{\sum_{k=1}^{n} exp\left(f_k G_T(u_k)\right)}$$

I change the broker's incentive structure by imposing that

$$\tilde{\rho}_{j,T} = \begin{cases} 1 \text{ if } u_j > u_l \, \forall l \in \mathcal{J}_{-j} \\ 0 \text{ otherwise} \end{cases}$$

Thus in a given market (defined in terms of the underlying equity and month), brokers must show the best available product in that market. For example, if a consumer is searching for a reverse convertible linked to Apple, the broker must show the client the best available Apple linked product in that month. The distribution of indirect utilities observed by consumers, $H_T(\cdot)$, is then generated by aggregating up the best products across each market. Note that even though consumers are always shown the best product in a given market, a consumer may still elect to continue searching across other markets. It is possible that the best product in a given market does not exceed the consumers reservation utility strategy.[31]

Table 7 displays the average change in search expenditures and consumer risk-adjusted returns under the new policy. On average, search expenditures for unsophisticated and sophisticated investors decline by 0.21% points and 0.01% points. These declines in search costs represent real increases in total economic surplus. Consumers capture most of the increase in surplus as consumer risk-adjusted returns for high and low types increase by 0.81% points and 0.01% points respectively. The average risk free adjusted rate over the period studied was 0.60%. Consequently these rep-

[31] As discussed in the preceding section, I assume for the empirical analysis that all consumers adopt the same reservation utility strategy across all of the observed markets.

resent relatively large gains in risk-adjusted returns. Just as with the preceding section, 8.1, this represents a partial equilibrium analysis in that the characteristics of available products are held fixed. Forcing brokers to always show the best available product in this manner would eliminate the market power currently held by brokers and issuers. For the same reasons described in Section 8.1, the estimates in Table 7 thus reflect a lower bound on the gain in expected consumer returns. As issuers optimally respond to the new broker policy, Nash Bertrand coupon/price competition among issuers will drive them to create the best possible product where issuers earn zero profits.

TABLE 7: ECONOMIC IMPACT OF BROKER INCENTIVES

	Consumer Type		
	Unsophisticated	Sophisticated	Average Consumer
Avg. Change in Search Expenditure	-0.21%	-0.0.01%	-0.21%
Avg. Change in Expected Return	0.81%	0.01%	0.80%

Table 7 Notes: Table 7 displays the hypothetical gains to total and consumer surplus if brokers were forced to always show the best product available in a market.

8.3 Financial Regulations

Across the globe regulators are moving towards addressing the asymmetry between broker and consumer incentives. Australia, the United Kingdom, India, Norway, Finland, Denmark and the Netherlands all recently placed bans on commissions in the financial service industry.[32] With the Dodd-Frank Act, US regulators are moving in a similar direction. As part of the Dodd-Frank Act, US regulators may soon require brokers to act as fiduciaries for their clients which would obligate brokers to act in the best financial interests of their clients.

One way of implementing fiduciary duty would be to force brokers to always show the best available product in each market as discussed in the previous section (8.2). The results displayed in Table 7 suggest that consumer returns would increase by 0.80% on a risk-adjusted basis. Perhaps

[32]See King & Wood Mallesons publication "Australia is Not Alone in Banning Commissions in Financial Services." http://www.mallesons.com/publications/marketAlerts/2012/regulator-october-2012/Pages/Australia-is-not-alone-in-banning-commissions-in-financial-services.aspx

even more importantly, total surplus would increase by 0.21% points relative to the total amount invested. Relative to the average profits earned in the financial sector,[33] a 0.21% point gain represents a substantial increase in surplus.

The analysis has a few important caveats. In the model, brokers service their client base by helping clients sequentially search over the product space. In practice, brokers may also impact the total number of investors in the market. If brokers were to leave the market in the event of new financial regulations, the total provision of financial services may also decline. Also, high fees might be justified for clients that are expensive for brokers to service (i.e., clients that require additional education or hand-holding).

Although current regulations target the incentives of brokers, the results in Tables 6 and 7 suggest that addressing search costs may be more beneficial. In the context of the model, if consumers had zero search costs, the broker incentive problem becomes irrelevant. Current SEC regulations restrict the marketing of financial securities directly to consumers, which could harm consumer welfare. The results suggest such regulations could be counterproductive if they raise consumer search costs.

9 Conclusion

Economists and regulators have long been interested in the observed price dispersion in financial products. Does such price dispersion imply consumers are overpaying for investments? Using a new data set I find evidence that consumers frequently purchase products with dominated payoff structures. What's even more alarming is that when both a superior and dominated product are available, consumers are more likely to end up with the latter.

Previous research has pointed to consumer search as the mechanism supporting price heterogeneity and potentially dominated financial products. Consumer search helps explain why consumers buy dominated products, but a standard search model cannot explain why consumers are more likely to purchase the dominated product over the superior product. I argue that consumers are more likely to purchase dominated products because the product fee structure incentivizes brokers to sell dominated products; hence, the incentives of brokers differ from the incentives of consumers. The empirical evidence verifies the incentive asymmetry. All else equal, consumers are more likely

[33] For example, Hansen et al. (2014) find the average profitability of bank deposits in the US is approximately 2%.

to buy products with higher fees. And similarly, all else equal, products with higher fees have lower payoffs.

The finding that consumers frequently overpay for investments and the finding that the incentives of brokers do not align with consumers are likely not unique to the reverse convertible industry. This paper focuses on reverse convertibles because some features of the reverse convertible market make identifying dominated products and the incentives of brokers easier. I find little reason to believe that search and broker incentives do not play important roles in other financial markets. A vast literature discusses price heterogeneity in financial markets, which suggests consumers might be overpaying for investments in other product markets (Hortaçsu and Syverson 2004, Gurun et al. 2013, and Green et al. 2007). Similarly, previous work, such as Livingston and O'Neal (1996), Mahoney (2004), Bergsteresser et al. (2009) and Christoffersen et al. (2013) details the potential conflict of interest arising in the mutual fund industry. The presence of dominated products and the broker/consumer incentive asymmetry prevalent in the market for reverse convertibles is more likely to be closer to the rule rather than the exception in financial markets.

References

Anagol, Santosh, Shawn Cole and Shayak Sakar. (2012). "Understanding the Incentives of Commissions Motivated Agents: Theory and Evidence from the Indian Life Insurance Market." Working Paper.

Barber, Brad M., Terrance Odean, and Lu Zheng. (2005). "Out of Sight, Out of Mind: The Effects of Expenses on Mutual Fund Flows." *The Journal of Business,* 78(6): 2095-2120.

Barseghyan, Levon, Francesca Molinari, Ted O'Donoghue and Josh Teitelbaum. (2013). "The Nature of Risk Preferences: Evidence from Insurance Choice." *American Economic Review*, 103(6): 2499-2529

Bergstresser, Daniel, John M. Chalmers, and Peter Tufano. (2009). "Assessing the Costs and Benefits of Brokers in the Mutual Fund Industry." *Review of Financial Studies*, 22(10): 4129-4156.

Black, Fischer and Robert C. Scholes. (1973). "The Pricing of options and Corporate Liabilities." *Journal of Political Economy*, 81: 637-654

Bolton, Peter, Xavier Freixas and Joel Shapiro. (2007). "Conflicts of Interest, Information Provision, and Competition in the Financial Services Industry." *Journal of Financial Economics*, 85(2): 297-330.

Brown, Jeffrey R., and Austan Goolsbee. (2002). "Does the Internet Make Markets More Competitive? Evidence from the Life Insurance Industry." *Journal of Political Economy*, 110(3): 481-507.

Carlin, Bruce I. (2009). "Strategic Price Complexity in Retail Financial Markets." *Journal of Financial Economics,* 91(3): 278-287.

Carlin, Bruce I., and Gustavo Manso. (2011). "Obfuscation, Learning, and the Evolution of Investor Sophistication." *Review of Financial Studies*, 24(3): 754-785.

Célérier, Claire and Boris Vallée. (2014). "What Drives Financial Complexity? A Look into the Retail Market for Structured Products." Working Paper.

Choi, James J., David Laibson, and Brigitte C. Madrian. (2010). "Why Does the Law of One Price Fail? An Experiment on Index Mutual Funds." *Review of Financial Studies*, 23(4): 1405-1432.

Christoffersen, Susan EK, Richard Evans, and David K. Musto. (2013). "What Do Consumers' Fund Flows Maximize? Evidence from their Brokers' Incentives." *The Journal of Finance,* 68(1): 201-235.

Christoffersen, Susan EK, and David K. Musto. (2002). "Demand Curves and the Pricing of Money Management." *Review of Financial Studies,* 15(5): 1499-1524.

Cronqvist, Henrik. (2006). "Advertising and Portfolio Choice." Working Paper.

Cronqvist, Henrik, and Richard H. Thaler. (2004) "Design Choices in Privatized Social-Security Systems: Learning from the Swedish Experience." *American Economic Review,* 94(2): 424-428.

Duarte, Fabian, and Justine S. Hastings. (2012). "Fettered Consumers and Sophisticated Firms: Evidence from Mexico's Privatized Social Security Market." NBER Working Paper No. w18582.

Elton, Edwin J., Martin J. Gruber, and Jeffrey A. Busse. (2004). "Are Investors Rational? Choices Among Index Funds." *The Journal of Finance*, 59(1): 261-288.

Gabaix, Xavier, and David Laibson. (2006) "Shrouded Attributes, Consumer Myopia, and Information Suppression in Competitive Markets." *The Quarterly Journal of Economics*, 121(1): 505-540.

Gil-Bazo, Javier and Pablo Ruiz-Verdú. (2009). "The Relation between Price and Performance in the Mutual Fund Industry." *The Journal of Finance*, 64(5): 2153-2183.

Green, Richard C. Burton Hollifield and Norman Shürhoff. (2007). "Dealer Intermediation and Price Behavior in the Aftermarket for New Bond Issues." *Journal of Financial Economics,* 86: 643-682.

Guran, Umit G., Gregor Matvos and Amit Seru. (2013). "Advertising Expensive Mortgages." NBER Working Paper No. w18910.

Hansen, Samuel G., Andrei Shleifer, Jeremy C. Stein and Robert W. Vishney. (2014). "Banks as Patient Fixed-Income Investors." NBER Working Paper No. 20288.

Hastings, Justine S., Ali Hortaçsu, and Chad Syverson. (2013). "Advertising and Competition in Privatized Social Security: The Case of Mexico." NBER Working Paper No. w18881.

Hastings, Justine S., and Lydia Tejeda-Ashton. (2008). "Financial Literacy, Information, and Demand Elasticity: Survey and Experimental Evidence from Mexico." NBER Working Paper No. w14538.

Haug, Espen Gaarder. (2007). *The Complete Guide to Option Pricing Formulas* (Vol. 2). New York: McGraw-Hill.

Heckman, James, and Burton Singer. (1984). "A Method for Minimizing the Impact of Distributional Assumptions in Econometric Models for Duration data." *Econometrica*, 52(2): 271-320.

Henderson, Brian J., and Neil D. Pearson. (2011). "The Dark Side of Financial Innovation: A Case Study of the Pricing of a Retail Financial Product." *Journal of Financial Economics,* 100(2): 227-247.

Hong, Han, and Matthew Shum. (2006). "Using Price Distributions to Estimate Search Costs." *The RAND Journal of Economics*, 37(2): 257-275.

Hortaçsu, Ali., and Syverson, Chad. (2004). "Product Differentiation, Search Costs, and Competition in the Mutual Fund Industry: A Case Study of S&P 500 Index Funds." *The Quarterly Journal of Economics,* 119(2): 403-456.

Jain, Prem C., and Joanna Shuang Wu. (2000). "Truth in mutual fund advertising: Evidence on future performance and fund flows." *The Journal of Finance*, 55(2): 937-958.

Khorana, Ajay and Henri Servaes. (2012). "What Drives Market Share in the Mutual Fund Industry." *Review of Finance,* 16: 81-113

Livingston, Miles and Edward S. O'Neal. (1996). "Mutual Fund Brokerage Commissions." *Journal of Financial Research*, 19(2): 273-292

Mahoney, Paul G. (2004). "Manager-Investor Conflicts in Mutual Funds." *The Journal of Economic Perspectives*, 18(2): 161-182.

Massa, Massimo. (2000). "Why So Many Mutual Funds?" Working Paper, INSEAD.

McCall, John Joseph. (1970) "Economics of Information and Job Search." *The Quarterly Journal of Economics,* 113-126.

Reiner, Eric and Mark Rubinstein. (1991a). "Breaking Down the Barriers." *Risk Magazine*, 4(8).

Reiner, Eric and Mark Rubinstein. (1991b). "Unscrambling the Binary Code." *Risk Magazine*, 4(9).

Rogerson, Richard, Robert Shimer and Randall Wright. (2005). "Search-Theoretic Models of the Labor Market: A Survey." *Journal of Economic Literature,* 43: 959-988.

Silverman, B.W. (1986). *Density Estimation.* London: Chapman and Hall.

Sirri, Erik R. and Peter Tufano. (1998). "Costly Search and Mutual Fund Flows." *Journal of Finance*, 53(5): 1589-1622.

Stahl, Dale O. (1989). "Oligopolistic Pricing with Sequential Consumer Search." *The American Economic Review*, 79(4): 700-712.

Stoimenov, Pavel A., and Sascha Wilkens. (2005). "Are Structured Products 'Fairly' Priced? An Analysis of the German Market for Equity-linked Instruments." *Journal of Banking & Finance*, 29(12): 2971-2993.

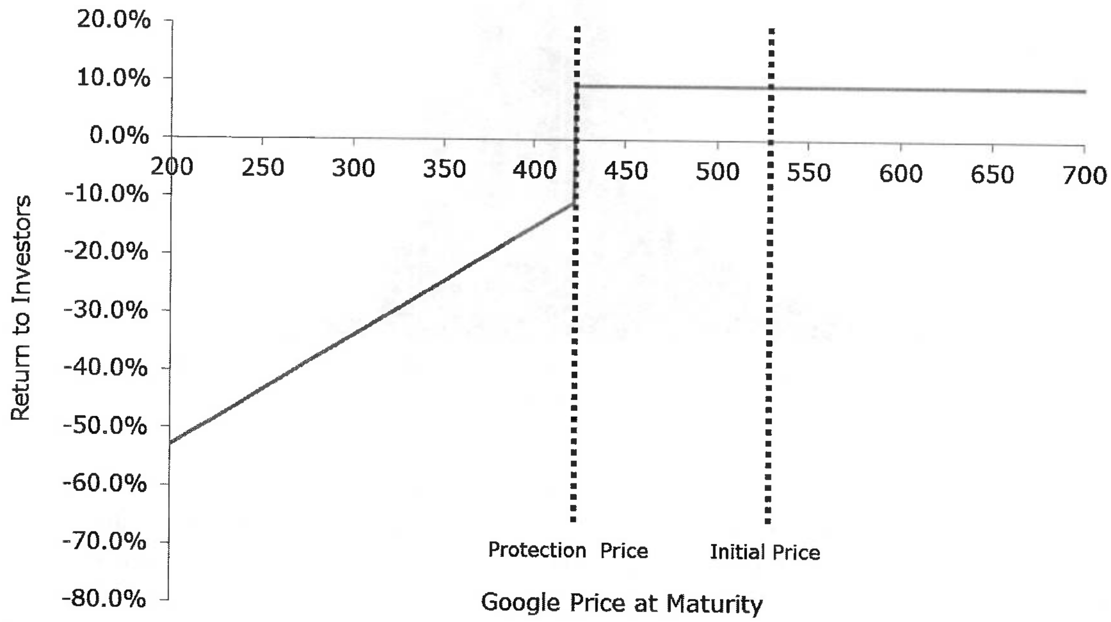
Szymanowska, Marta, Jenke Ter Horst, and Chris Veld. (2009). "Reverse Convertible Bonds Analyzed." *Journal of Futures Markets*, 29(10): 895-919.

Wahal, Sunil and Albert (Yan) Wang. (2011). "Competition Among Mutual Funds." *Journal of Finance*, 99: 40-59.

Woodward, Susan E., and Robert E. Hall. (2012). "Diagnosing Consumer Confusion and Sub-optimal Shopping Effort: Theory and Mortgage-Market Evidence." *American Economic Review,* 102(7): 3249-3276.

Appendix

FIGURE A-1: REVERSE CONVERTIBLE EXAMPLE (SINGLE OBSERVATION)



hoose

to

Figure A-1 Notes: The figure displays the return to investors for a one year reverse convertible bond linked to the price of Google Inc. that was issued by UBS (CUSIP 90268F112). The reverse convertible pays a monthly coupon of 9.25%. If at maturity the price of Google closes above the protection price (convertible trigger price) of $422.63 (80% of the initial price), investors will receive 100% of the principal at maturity earning a return of 9.25%. If the share price of Google Inc. closes below $422.63, the issuer will pay the bondholder 1.89 shares of Google Inc. per $1,000 invested ($1,000/Initial Price) rather than 100% of the principal amount invested. The above figure displays the final return to investors based on the price of Google Inc. at maturity.

FIGURE A-2: DISPERSION IN RISK ADJUSTED RETURNS

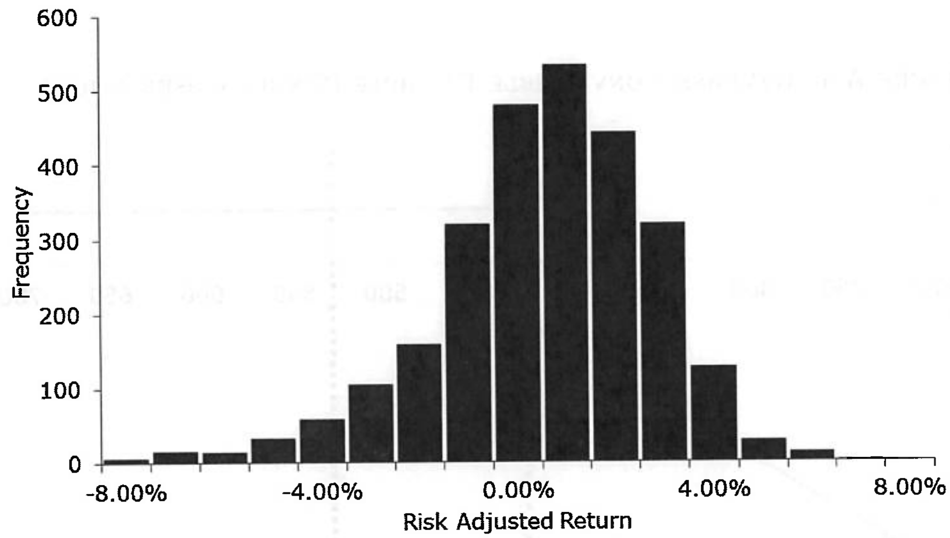


Figure A-2 Notes: The figure displays dispersion in reverse convertible risk-adjusted returns. I calcute the risk adjsted return of each reverse convertible as the present value of coupon payments (assuming monthly coupons discounted using the one year swap rate) minus the implied option premium, the issuer CDS spread and the one year risk free rate (as measured using the one year swap rate). I normalized risk-adjusted returns such that the average return is zero.

Table A-1: Structural Estimation Results

Variables	(1)	(2)
Coupon (α)	1	1
Option Premium (γ^{Delta})	-0.83***	-1.17***
	(0.07)	(0.0135)
CDS Spread (γ^{CDS})	-4.55***	-4.05***
	(0.96)	(0.17)
Scaling Parameter(θ)	49.27***	36.49***
	(15.22)	(13.45)
π_H		1.56%***
		(0.25%)
Heterogeneous Agents		X
Observations	1,227	1,227
Number of Markets	423	423

Table A-1 Notes: Table A-1 displays the maximum likelihood estimation results for the fully specified model. Each observation is weighted by the market size. Standard errors are calculated using the observed Fisher Information Matrix., **, *** indicate significance at the 10%, 5% and 1% level.*

34

January 2013, Number 13-1

CENTER for RETIREMENT RESEARCH at BOSTON COLLEGE

HOW DO EMPLOYERS' 401(k) MUTUAL FUND SELECTIONS AFFECT PERFORMANCE?

By Edwin J. Elton, Martin J. Gruber, and Christopher R. Blake*

Introduction

Defined contribution plans, predominantly 401(k)s, are the primary source of personal retirement savings for American workers, making the investment decisions within these accounts a salient policy concern.[1] These decisions are a result of two separate actions: the mutual fund options selected by the employer's plan administrator and the specific funds chosen by the participant.

While considerable research has examined 401(k) participant decisions in isolation, surprisingly little attention has been focused on the choices made by plan administrators. The administrator's role is clearly influential, particularly if, as indicated by prior research, 401(k) participants themselves do not make good choices. This *brief*, based on a prior study, addresses this research gap by focusing on the fund choices of 401(k) plan administrators and participants' reactions to these choices.[2]

The discussion proceeds as follows. The first section reviews existing research on 401(k) investment decisions. The second section explains the data and the metric used to analyze how employer and employee fund choices affect investment performance. The third section explores how well plan administrators do in choosing mutual funds. The fourth section assesses how well participants do. The fifth section concludes that employers select mutual funds that perform better than comparable, randomly selected, funds but worse than passive index funds, and participants do not add any value through their own decisions.

401(k) Investment Decisions: What We Know

Due to the growing influence of 401(k)s, researchers have examined numerous aspects of the investment choices made by plan participants. Virtually all the findings suggest that the individual investor does not make very good decisions. One study found that participants restrict their investing to three or four mutual funds – regardless of how many funds their employer offers.[3] Other research finds that employees simply divide their savings evenly among the number of funds (N) their employers offer – a strategy known as the 1/N Rule.[4] Other studies examining asset allocation find that plan participants infrequently adjust their allocations; that their ages and cohorts influence their stock allocations; and that they over-invest in their employer's stock, which reduces diversification.[5] In short, the consistent message is that participants often make poor choices.

* Edwin J. Elton and Martin J. Gruber are both professors emeriti and scholars in residence at New York University (NYU)'s Stern School of Business. Christopher R. Blake is a professor of finance and Joseph S. Keating, SJ, Distinguished Professor of Business at Fordham University. This *brief* is based on Elton, Gruber, and Blake (2007).

All of these previous studies examined participant decisions only. But plan administrators also have a major role as they select a limited menu of mutual funds to offer participants from the large number of available funds. One study that did examine administrator choices found that about one half of plans do not provide sufficient categories of investments to their participants.[6] This *brief* builds on this study by examining whether, given the categories of investments offered, the fund choices selected by plan administrators are good investments per se, and how participants react to the choices.[7]

Data and Methodology

The main data source for both the employer and employee analyses is the 11-K report that an employer files annually with the Securities and Exchange Commission if its 401(k) plan offers the company's own stock as an investment option.[8] The period of analysis covers 1994-1999; data after 1999 were unavailable due to a change in the Securities and Exchange Commission's electronic filing requirements.

Mutual fund selections and performance are analyzed for each plan in the sample; plans are eliminated if they provide data only by broad investment categories such as stocks, bonds, or a specific mutual fund family. This process leaves a sample of 43 plans with individual mutual fund data and an average asset size of $310 million.[9]

Three other types of data are also required. Monthly investment returns for individual mutual funds are from the University of Chicago's Center for Research in Securities Prices (CRSP). Monthly returns for indexes, which are used as benchmarks for performance comparisons, are obtained from CRSP, Morningstar, and a private website. Finally, when a risk-free interest rate is required in the analysis, the yield on 30-day U.S. Treasury bills is used.

The key metric used to gauge investment performance is "alpha," which is the rate of return above or below what would have been earned on a passive portfolio of indexes with the same risk profile. Alpha can be computed for each mutual fund offered and these fund-specific alphas can then be combined to compute an alpha for each employer's 401(k) plan. A positive alpha indicates that the mutual funds in a plan outperformed their benchmark indexes; a negative alpha indicates their performance did not keep pace.[10] Alpha, on average, is negative, because "active" funds managed by stock pickers generally underperform their relevant market indexes.[11] "Passive" mutual funds typically have a negative alpha as well due simply to the fees charged to manage the fund.

The analysis summarized below reports two measures: 1) an alpha for the combined funds in each 401(k) plan relative to a passive portfolio of indexes; and 2) a "differential alpha," which is the difference between the alpha for each 401(k) plan and the average alpha for a randomly selected sample of similar funds.[12]

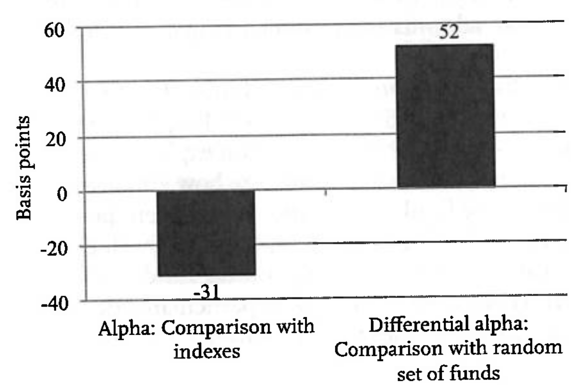
Performance of Plan Administrators

The performance of administrators is evaluated in two ways: 1) by how well each plan's mutual funds do compared to the benchmark indexes (alpha) and to a random sample of similar funds (differential alpha); and 2) by how well funds that were added or dropped perform both before and after the switch.[13]

How Well Do Funds Perform?

The results for the sample plans show that the average alpha over three years of investment performance is -31 basis points annually (see Figure 1). The negative alpha, as expected, confirms that the plans' performance falls below the performance of comparable indexes.[14] The size of this negative alpha is larger than normal expenses for low-cost index funds, suggesting that performance would be improved if passive funds had been substituted for the active funds that were selected.

Figure 1. Performance of Sample 401(k) Plans Based on Alpha and Differential Alpha, in Basis Points Per Year



Note: Results assume equal weighting of each fund within an employer's 401(k) plan.
Source: Elton, Gruber, and Blake (2007).

The average differential alpha for the sample 401(k) plans, however, was +52 basis points annually. This result shows that plan administrators, overall, chose mutual funds that outperformed the randomly selected set of funds by about one-half of 1 percentage point annually.

Lower investment fees are a large part of the explanation for the superior performance of the employer selections compared to the random set of funds. Lower fees, by definition, improve returns by leaving more money in the investor's account. The fees in the employer-selected mutual funds were 23 basis points per year lower than the fees for the random set of funds, accounting for almost half of plan administrators' superior results.

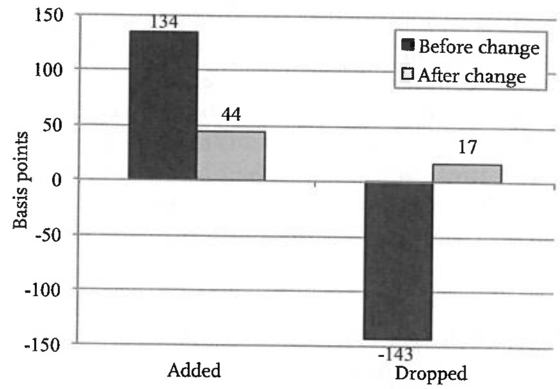
Do Fund Changes Improve Performance?

401(k) investment performance can also be influenced by changes in mutual fund offerings over time. During the period analyzed, the employers in the sample added 215 mutual funds and dropped 45 funds. Many of the additions seem to be motivated by a desire to add a new type of fund, as over half were selected from an investment category not held by the plan at the time of the addition.

The analysis looked at the performance of the added and dropped funds for three years *before* the change was made and three years *after* the change. Not surprisingly, newly added funds outperformed randomly selected funds before the change was made: the differential alpha of the added funds is +134 basis points annually for three years prior to being added to the sample's 401(k) plans. In contrast, before the dropped funds were dropped, they under-performed the random funds by -143 basis points annually. Thus, the added funds outperformed the dropped funds by a total of 277 basis points annually prior to when the changes were made (see Figure 2).

Interestingly, though, this performance bonus essentially disappeared *after* the fund changes were made as the added funds did worse while the dropped funds did better. The differential alphas after the changes are +44 basis points for the added funds and +17 for the dropped funds, and the difference between them is not statistically significantly different from zero. This finding suggests that plan managers were chasing returns, but their efforts to tinker with their fund selections had essentially no impact on overall performance. The outcome underscores the traditional investor's caveat that "past performance does not predict future returns."

Figure 2. Performance of Added Funds and Dropped Funds Based on Differential Alpha, in Basis Points Per Year, Before and After Change



Note: The gap between the added and dropped funds after the changes were made (indicated by the gray bars) is not statistically significant.
Source: Elton, Gruber, and Blake (2007).

Performance of Plan Participants

This section turns to the performance of 401(k) participants to see whether their behavior is consistent with that depicted in the existing literature and to assess whether they add value to the decisions made by plan administrators. The first exercise evaluates whether participants rebalance their portfolio in response to market fluctuations or, instead, chase returns. The second exercise compares the participants' investment strategies, at an aggregate level for each plan, to naïve investment strategies.

Do Participants Chase Returns?

Three factors influence asset allocation: annual returns, participant contributions,[15] and participant transfers. For all sample plans, the median change in the percent of assets allocated to particular investments over all the years analyzed is 3.8 percentage points for investment returns, 1.6 percentage points for participant contributions, and 3.1 percentage points for participant transfers. These numbers indicate how the distribution of assets between mutual funds changes over time. While investment performance has the largest impact on the weightings, participants also have a significant impact when they alter their contributions or transfer assets.

The next step is to determine whether participants' actions magnify or offset the change in allocations caused by investment returns. A regression analysis relates the combined effect of participants' contributions and transfers to the effect of returns for each of the sample plans. The results show that participants' contributions and transfers magnify the change in allocations caused by returns by 57 percent. That is, participants shift their assets toward the best-performing funds and decrease their holdings in the funds that do not perform as well, causing the fund allocations to diverge further from the plans' initial weightings.

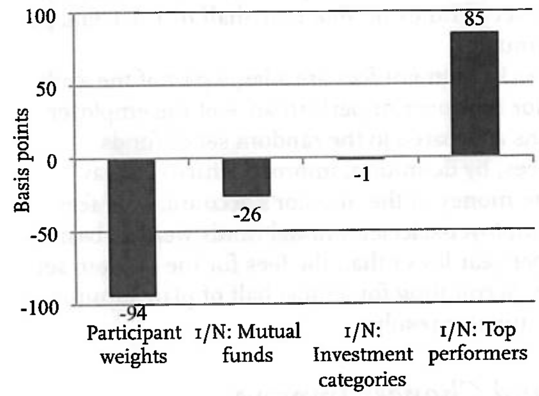
Do Participants Outperform Naive Investment Strategies?

The final analysis examines whether participants' decisions, in aggregate, improve or worsen their 401(k) investment performance. Participants' impact on performance is gauged through a comparison with what their returns would have been if they had instead adopted the simple 1/N Rule, in which investors spread their assets evenly across all of the funds.

The alpha measure is calculated for participants' returns based on their actual investment choices. This measure is then compared to three versions of the 1/N rule: 1) the simple 1/N Rule in which equal allocations are made to each fund; 2) a 1/N Rule in which equal allocations are made to each *investment category*; and 3) a 1/N Rule using only mutual funds with investment performance that fell in the top half of all the funds available.

The results in Figure 3 show that the participants' actual selections performed no better than any of the 1/N strategies. In fact, the participants' results were lower in all cases, though only the difference with the "top performers" strategy was statistically significant at the 5-percent level. These results suggest that participants in aggregate do not add value to the investment performance of their 401(k) through their own decisions, underscoring the importance of the choices made by plan administrators.

FIGURE 3. PERFORMANCE USING PARTICIPANTS' OWN FUND WEIGHTINGS AND THREE 1/N RULES BASED ON ALPHA, IN BASIS POINTS PER YEAR



Note: Estimates are for one-year performance.
Source: Elton, Gruber, and Blake (2007).

Conclusion

The mutual funds that 401(k) administrators select achieve investment returns that are worse than comparable indexes but superior to the returns of comparable, randomly selected funds. A significant part of this latter result is explained by choosing funds that charge lower fees. When making changes to a plan's funds, administrators chase returns and do not end up improving investment performance.

Like their employers, 401(k) plan participants also tend to chase returns, transferring assets into higher-performing funds, rather than rebalancing to restore their original asset allocations. And their investment performance is no better than they would have achieved using variations on the 1/N rule to allocate assets among funds.

Endnotes

1 According to data from the Federal Reserve's *Flow of Funds*, defined contribution plans held over $4 trillion in 2012, with an additional $5 trillion in individual retirement accounts (IRAs) that mostly represents rollovers from defined contribution plans.

2 Elton, Gruber, and Blake (2007).

3 Huberman and Jiang (2006).

4 Benartzi and Thaler (2001).

5 For frequency of allocation changes, see Mitchell et al. (2005) and Madrian and Shea (2001). For impact of employee ages on allocations, see Agnew and Balduzzi (2004). For investment in company stock, see Huberman and Jiang (2006).

6 Elton, Gruber, and Blake (2006).

7 A recently published study by Brown and Harlow (2012) also examined plan administrator choices. It reinforces two of the key findings in the study summarized in this *brief* (Elton, Gruber, and Blake 2007), specifically that the options employers offer to their plan participants do not outperform index funds and do outperform actively-managed mutual funds.

8 One advantage of 11-K filings is that a number of years of data are available to show participant behavior and plans offered by fund families. The disadvantage is that only aggregate – rather than individual – 401(k) participant impacts can be examined.

9 This amount compares with the average $411 million asset size for data used by Liang and Weisbenner (2002).

10 The three-year alpha calculations begin with the date on each employer's 11-K report. For the index benchmarks, alphas are calculated over the three years following the end of each fund's fiscal year.

11 See, for example, Blake, Elton, and Gruber (1993), Elton, Gruber, and Blake (1996), and Grinblatt and Titman (1989).

12 For the purposes of this analysis, a "similar" fund is one in the same Investment Company Data, Inc. (ICDI) investment-objective category and of similar size.

13 This *brief* covers only selected portions of the full analysis presented in Elton, Gruber, and Blake (2007).

14 This *brief* presents results on the alpha and differential alpha for three years of investment returns and assumes that each mutual fund in the 401(k) plan has an equal weight. Elton, Gruber, and Blake (2007) also estimate the differential alpha for a one-year period and for an alternative weighting assumption that weights each fund in a plan according to participants' actual allocations to each fund. The results for these alternative assumptions are broadly similar to those presented here.

15 Employers also make contributions to participants' 401(k) plans. The results for participant contributions are reported separately here as the focus is on participant decisions, and the allocation of employer contributions is sometimes determined by the plan itself rather than by participants. For results that include employer contributions, see Elton, Gruber, and Blake (2007).

References

Agnew, Julie and Pierluigi Balduzzi. 2004. "Large, Small, International: Equity Portfolio Choices in a Large 401(k) Plan." Chestnut Hill, MA: Center for Retirement Research at Boston College.

Bernatzi, Shlomo and Richard Thaler. 2001. "Naïve Diversification Strategies in Retirement Savings Plans." *American Economic Review* 91(1): 78-98.

Brown, Keith C. and W. Van Harlow. 2012. "How Good Are the Investment Options Provided by Defined Contribution Plan Sponsors?" *International Journal of Portfolio Analysis and Management* (1): No. 1.

Blake, Christopher R., Edwin J. Elton, and Martin J. Gruber. 1993. "The Performance of Bond Mutual Funds?" *Journal of Business* 66(3): 371-403.

Elton, Edwin J., Martin J. Gruber, and Christopher R. Blake. 2007. "Participant Reaction and the Performance of Funds offered by 401(k) Plans." *Journal of Financial Intermediation* (16): 249-271.

Elton, Edwin J., Martin J. Gruber, and Christopher R. Blake. 2006. "The Adequacy of Investment Choices Offered by 401(k) Plans." *Journal of Public Economics* (90): 1299–1314.

Elton, Edwin J., Martin J. Gruber, and Christopher R. Blake. 1996. "The Persistence of Risk-Adjusted Mutual Fund Performance." *Journal of Business* 69(2): 133-157.

Grinblatt, Mark and Sheridan Titman. 1989. "Mutual Fund Performance: An Analysis of Quarterly Portfolio Holdings." *Journal of Business* 62(3): 393-416.

Huberman, Gur and Wei Jiang. 2006. "Offering Versus Choice in 401(k) Plans: Equity Exposure and Number of Funds." *Journal of Finance* (61): 763-801.

Liang, J. Nellie and Scott Weisbenner. 2002. "Investor Behavior and the Purchase of Company Stock in 401(k) Plans: The Importance of Plan Design." Unpublished manuscript. University of Illinois.

Madrian, Brigitte and Dennis F. Shea. 2001. "The Power of Suggestion: Inertia in 401(k) Participants' Savings." *Quarterly Journal of Economics* 116 (4): 1149-1187.

Mitchell, Olivia, Gary Mottola, Stephen Utkus, and Takeshi Yamaguchi. 2005. "The Inattentive Participant: Portfolio Trading Behavior in 401(k) Plans." Working Paper 2006-2. Philadelphia, PA: Pension Research Council.

U.S. Board of Governors of the Federal Reserve System. 2012. *Flow of Funds Accounts of the United States*. Washington DC.

CENTER *for*
RETIREMENT
RESEARCH
at BOSTON COLLEGE

About the Center

The Center for Retirement Research at Boston College was established in 1998 through a grant from the Social Security Administration. The Center's mission is to produce first-class research and educational tools 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. Since its inception, the Center has established a reputation as an authoritative source of information on all major aspects of the retirement income debate.

Affiliated Institutions

The Brookings Institution
Massachusetts Institute of Technology
Syracuse University
Urban Institute

Contact Information

Center for Retirement Research
Boston College
Hovey House
140 Commonwealth Avenue
Chestnut Hill, MA 02467-3808
Phone: (617) 552-1762
Fax: (617) 552-0191
E-mail: crr@bc.edu
Website: http://crr.bc.edu

The Center for Retirement Research thanks AARP, Advisory Research, Inc. (an affiliate of Piper Jaffray & Co.), Citigroup, Invesco℠, Mercer, National Reverse Mortgage Lenders Association, Prudential Financial, State Street, TIAA-CREF Institute, T. Rowe Price, and USAA for support of this project.

The research reported herein was supported by the Center's Partnership Program. The findings and conclusions expressed are solely those of the authors and do not represent the views or policy of the partners or the Center for Retirement Research at Boston College.

35

ELSEVIER

Journal of Public Economics 90 (2006) 1299–1314

JOURNAL OF PUBLIC ECONOMICS

www.elsevier.com/locate/econbase

The adequacy of investment choices offered by 401(k) plans

Edwin J. Elton [a], Martin J. Gruber [a,*], Christopher R. Blake [b]

[a] *Department of Finance, Stern School of Business, 44 West Fourth Street New York, NY 10012, United States*
[b] *Fordham University, United States*

Received 2 January 2005; received in revised form 27 July 2005; accepted 9 August 2005
Available online 19 September 2005

Abstract

The choices made by 401(k) participants are the product of two different decisions: what is offered and what is chosen. While there have been a number of studies of the decisions made by participants in 401(k) plans, there have been no studies of the adequacy of the full set of choices offered to 401(k) participants. This paper analyzes the adequacy and characteristics of the choices offered to 401(k)-plan participants for over 400 plans. We find that only 53% of the plans offer an adequate set of options and that over a 20-year period offering inadequate options makes a difference in terminal wealth of over 53%. We find that funds included in the plans are riskier, but have a slightly higher return, than the general population of funds in the same categories. However, we find that the return difference is roughly equal to the difference in expenses between funds selected by plans and randomly selected funds. We study the characteristics of plans that are associated with adequate investment choices, including an analysis of the use of company stock, plan size, and the use of sophisticated strategies.

PACS: G23; G11
Keywords: 401k pension funds; Portfolio performance mutual funds; Spanning

The value of any 401(k) pension plan to any participant is determined by two decisions: the set of investment choices offered to the participant in a plan, and what the participant invests in from among those choices. These are two very different decisions. The first decision is made by the plan administrator; the second is made by the participant in the plan. There is a large amount

* Corresponding author. Tel.: +1 212 998 0333; fax: +1 201 670 4710.
E-mail address: mgruber@stern.nyu.edu (M.J. Gruber).

0047-2727/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.jpubeco.2005.08.001

of research on the participant's investment behavior, given the choices available to the participant.[1] However, none of this research examines the impact of the full set of investment choices offered in a plan on the ability of plan participants to construct desirable portfolios.[2]

Examining the choices offered to the participants in any plan is important, because an investor faced with an inappropriate set of choices cannot construct an efficient portfolio from among these choices no matter what weights he places on the various offerings. There are two ways that the choices offered can be inappropriate: offering an insufficient number and type of choices to allow the construction of desirable portfolios, and offering poor-performing investment choices of any given type.

For some participants, insufficient choices can be alleviated by holding assets outside their 401(k) plan. However, for the largest group of 401(k) participants, those earning between \$20,000 and \$70,000 annually, the participants have less than one month's income invested outside of their pension savings.[3] Furthermore, in 2001 more than 60% of participants in 401(k) plans had no other investments (stocks, bonds, etc.) outside of employer plans.[4] For twenty-eight percent of the participants, a spouse or partner also had a retirement plan. In addition, some of the participants also had a defined-benefit plan.[5] Nevertheless, a substantial fraction of 401(k) plan participants are solely dependent on their 401(k) plan offerings to fund retirement. While these participants represent a smaller percentage of wealth than they represent of the total number of participants, they represent the segment of the population that is most dependent on 401(k) plans for retirement.

What choices should a corporation offer to plan participants? For those participants for whom 401(k) investments are their sole financial assets, the corporation should offer a sufficient set of investment alternatives so that the investor could construct the same efficient frontier that he or she would obtain if there were choices from a reasonable set of alternatives. Investors who have other financial assets would not be hurt by such a strategy, so this strategy is dominant for all investors.

In this paper we find that only 53% of pension plans offer a set of choices that allows investors to construct an efficient frontier equivalent to one constructed from a reasonable set of alternative choices and that the loss because of this to investors is substantial. Second, we find that plans offer funds that have better performance than randomly selected funds but that the difference is about the same size as the expense difference between the funds they select and randomly selected funds. Third, we show that the portfolio of mutual funds offered by 401(k) plans is riskier than randomly selected mutual funds under realistic assumptions. Fourth, we show that the inclusion of company stock in a plan does not affect the desirability of the other choices offered to plan participants.

[1] Examples of this extensive literature are Benartzi and Thaler (2001), Liang and Weisbenner (2002), Huberman and Sengmuller (2004), Agnew and Balduzzi (2003), Ameriks and Zeldes (2004), Madrian and Shea (2001).

[2] The paper closest in spirit to this one is by Angus, Brown, Smith and Smith (2005), who examined the appropriateness of the choices offered by one 403(b) plan (TIAA-CREF). Other authors, for example Choi, Laibson, Madrian and Metrick (in press), have examined the effects of certain aspects of the choice sets offered by plans such as the inclusion of company stock or the use of default options.

[3] See Choi, Laibson, Madrian and Metrick (2004).

[4] See Investment Company Institute (2000).

[5] Whether a defined-benefit plan helps an inadequate set of offerings become adequate is unclear. The investor cannot choose the offerings in the defined-benefit plan. Furthermore, changes in the present value of the payments from the plan are a function of the discount rate and the participant's salary. How changes in the present value of the defined-benefit plan help an inadequate set of 401(k) choices is ambiguous.

This paper is divided into five sections. In the First section we discuss the data used. In the Second section we explore issues of how well the fund offerings span the efficient set. In that section, we not only examine statistical tests, but we also examine the economic significance (effect on participants' returns) of a failure to provide appropriate offerings. In the Third section we explore the characteristics (risk-adjusted return and risk) of the specific funds offered relative to the population of funds that could have been offered. In the Fourth section we examine the effect of offering company stock on plan risk and the efficient frontier. In Section five we examine whether other characteristics of the plans affect the appropriateness of the investment choices offered to plan participants. Finally, in Section six we summarize our results.

1. Data

Our data were provided by Moody's Investor Services. Moody's collects data by means of a survey of pension plans offered by both for-profit and non-profit entities (collected in 2002 with information through 2001). From this data set we selected all 401(k) plans (plans offered by for-profit entities) that employed publicly available mutual funds for participant choices. However, we did not exclude plans that offered, in addition to mutual funds, private money market funds, GICs and stable value funds and/or company stock. We were able to identify 680 401(k) plans for which the CRSP mutual fund database contained at least some data on each of the mutual funds offered in the plan. Of the 680 plans, 417 had at least five years of monthly total returns data in the CRSP database for every mutual fund they offered.[6] For each of these plans we collected data on the mutual funds offered, historical returns for each mutual fund, and the names and characteristics of the firms offering the plans.

Table 1 shows the number of distinct investment choices offered by the 680 plans mentioned above. The median number of 401(k) plan offerings is eight. Approximately 12% of the 401(k) plans offer four or fewer investment choices, and approximately 11% offer 13 or more investment alternatives. The median number of investment offerings we report is somewhat less than that reported by Huberman and Sengmuller (2004). Huberman and Sengmuller's data sample came from 401(k) plans administered by Vanguard. Many plans restrict their offerings to one fund family. Vanguard is one of the largest mutual fund families in terms of number of funds offered. Thus it is not surprising that plans managed by Vanguard offer more choices than would be observed in the population.[7] The Investment Company Institute (2000) sampled 1181 households and found that the median number of choices offered by the plans to which the households belonged was six. Compared to that study, the number we find is slightly larger. Considering the two studies together, the number of funds held by plans in our sample is consistent with what others have found.

Table 2 presents the percentage of plans that offer various types of investment choices to their participants using Investment Company Data, Inc. (ICDI) classifications. The most common investment choice (offered by 97.4% of plans) is a domestic equity fund. The next most common offering (86.8%) is an alternative such as a GIC or money market fund, where interest is intended to be the only source of return. Other common offerings fall in the following categories: domestic bond funds (71.5%), domestic mixed bond and stock funds (80.6%), and international

[6] When later we draw samples of mutual funds for comparison purposes, we use the same selection procedure, so that the comparisons are unbiased.

[7] When we examine plans that only use Vanguard funds, we get numbers similar to those reported by Huberman and Sengmuller.

Table 1
Percentages of 680 401(k) plans offering different numbers of investment choices (number of choices and percentages include all offerings by the plan, including mutual funds, stable value funds, GICs and company stock)

Number of investment choices	Percentage of plans
1	2.21%
2	2.35%
3	3.09%
4	4.85%
5	8.97%
6	12.06%
7	12.06%
8	13.82%
9	11.76%
10	9.85%
11	5.59%
12	2.21%
13	2.50%
14	1.91%
15	1.18%
16	1.03%
17 or more	4.56%

This table is based on 2001 survey data of 680 401(k) plans obtained from Moody's Investor Services.

Table 2
Types of investment choices offered in 680 401(k) plans

Category	ICDI classification	Percentage of plans offering investment choice
Interest only	Money market fund	57.35%
	Stable value fund	24.85%
	GIC	14.71%
	At least one interest-only fund	86.76%
Domestic equity	Aggressive growth fund	55.44%
	Growth and income fund; equity index	80.00%
	Long-term growth fund	82.94%
	Sector fund	6.47%
	Total return fund; equity value fund	21.62%
	Utilities fund	0.59%
	At least one domestic equity fund	97.35%
Domestic bonds	Quality bond fund	54.85%
	High-yield bond fund	4.71%
	Government mortgage fund	12.79%
	Government securities fund	12.35%
	At least one domestic bond fund	71.47%
Domestic mixed	Balanced fund	73.68%
	Income fund	23.68%
	At least one domestic mixed fund	80.59%
International	Global bond fund	9.12%
	Global equity fund	18.97%
	International equity fund	62.94%
	At least one international fund	75.15%
Company stock	Company stock	22.94%

This table is based on survey data of 680 401(K) plans obtained from Moody's Investor Services.

bond and/or stock funds (75.1%). The high percentage of 401(k) plans that offer international funds is surprising, given the much lower percentage international funds constitute of mutual funds publicly available to investors. Finally, 22.9% of the 401(k) plans offer company stock as an alternative for their participants.[8]

2. Adequacy of investment choices

In this section we examine the adequacy of the investment choices offered by 401(k) plans. In order to determine if 401(k) plans offer their participants appropriate investment choices, we need to hypothesize an adequate set of alternative investment choices. We construct the alternatives by drawing on the literature of financial economics, which discusses indexes that are necessary and sufficient to capture the relevant return characteristics for a range of investments.[9]

The indexes employed will now be described in greater detail. For common stocks, we classify by value versus growth and by size as advocated by Fama and French (1995). We classified size into two groups: small-mid-cap and large-cap. Each of these two groups was then further divided into value and growth. All four indexes were taken from Wilshire. We chose Wilshire indexes because there exist tradeable funds that attempt to match each of the indexes. For bonds, we combined a general bond index, including governments and corporates, and a mortgage-backed index. We also employed a high-yield index. This division is supported by Blake, Elton and Gruber (1993), who found this division was sufficient to capture differences in return across bond funds. In the combined bond index, we used the Lehman U.S. Government/Credit index and the Lehman Fixed-Rate Mortgage-Backed Securities index. We used the Credit Suisse First Boston High-Yield index for the high-yield bond index.[10] We also included the Salomon Non-U.S.-Dollar World Government Bond index for international bonds and the MSCI EAFE index for international stocks.

Since, in this study, returns on all mutual funds are computed after expenses, we deducted expenses from each of our indexes. For each of our indexes, we used the expense charge of the index fund (including exchange-traded funds) that most closely matched the index. If there were multiple index funds matching the index, we used the expense charge of the lowest cost fund.

We now examine whether the choices given investors allow construction of an efficient frontier similar to that obtained by the 8 indexes. To do this we use intersection tests.

2.1. Methodology

The purpose of the intersection test is to examine whether, given a riskless rate, a particular set of assets is sufficient to generate the efficient frontier or whether including (long or possibly short) members of a second set of assets would improve the efficient frontier at a statistically

[8] Across the 680 plans in our sample, a total of 56 specialized funds were offered. We noted that 33 of these funds were T. Rowe Price funds, suggesting that recommending inclusion of specialized funds may be part of T. Rowe Price's marketing strategy to 401(k) plans.

[9] In an earlier version of this paper we compared the commonly used industrial classifications of indexes with the indexes from financial economics and found that the indexes from the literature were superior in explaining returns. The results are available from the authors.

[10] Originally the indexes for mortgage-backed securities and small- and mid-cap stocks were included separately, but empirical tests showed there was no improvement in explaining returns by including them separately. Results are available from the authors.

significant level. In other words, given the best portfolio (linear combination) of the assets held by a plan, can adding one or more of the eight indexes to this portfolio statistically significantly improve the return for the portfolio at a given level of risk or do the original plan offerings span the eight indexes?

As DeRoon, Nijman and Werker (2001) have shown, intersection is a test of the impact of restricting the intercept (α) in the following time-series model:

$$R_{it} - R_{\mathrm{f}} = \alpha_i + \sum_{k=1}^{K} \beta_{ik}(R_{kt} - R_{\mathrm{f}}) + \varepsilon_{it} \tag{1}$$

Where

R_{it}	the return on one of the eight indexes described earlier in month t ($i=1,\ldots,8$);
R_{f}	the risk-free rate;
R_{kt}	the return on fund k in the plan in month t;
ε_{it}	the error term for index i in month t;
β_{ik}	the sensitivity of index i to plan fund k.

Investors can only hold 401(k) plan assets and additional mutual funds long. When short sales are not allowed, the right-hand side of Eq. (1) includes returns on only those funds in a plan that are held long in the optimal portfolio of plan funds.[11] Intersection occurs if, for all of the eight indexes jointly, the α_i are not statistically significantly positive, i.e., the restrictions are

$$\alpha_i \le 0 \ \forall \ i. \tag{2}$$

The logic behind the test can be easily understood.[12] The intercept is the additional risk-adjusted return that is available on each of the eight indexes that is not available on a linear combination of the plan assets. If short sales are forbidden, then only the addition of an asset with a positive alpha can improve the efficient frontier, by offering a higher return on an optimal portfolio for any given level of risk. Since most assets would have a positive or negative alpha, the test examines whether the shift in the efficient frontier associated with positive alphas is significant. Since adding assets to any portfolio has a probability of improving efficiency, it is important to employ a test of statistical significance. The test we employ is the intersection test, given a riskless rate and short sales not allowed, shown in DeRoon, Nijman and Werker (2001).[13]

2.2. Results

The results of the intersection tests are shown in Table 3. Recall that the intersection tests would not reject spanning if adding any of the eight indexes to the optimal portfolio of funds offered by any plan does not improve the efficient frontier at a statistically significant level. Plans where spanning is not rejected offer participants a sufficient set of choices. The majority of plans holding four or fewer funds do not offer a set of funds that span the eight indexes. For

[11] These assets can be easily identified by solving a quadratic programming problem for each plan.

[12] For the theory and application of spanning tests, see, for example, Beckaert and Urias (1996), Chen and Knez (1996), Hansen, Heaton and Luttmer (1995), Huberman and Kandel (1987) and Glen and Jorion (1993).

[13] The test employs the Wald test, recognizing that, with inequality constraints, the relevant comparison to determine statistical significance is based on a mixture of chi-square distributions.

Table 3
Sufficiency of plan investment choices in spanning eight indexes (short sales not allowed)

Number of investment choices in plan	Total number of plans	Number of plans that span (offer sufficient choices)
1	10	3
2	18	4
3	37	18
4	57	28
5	53	33
6	58	24
7	44	26
8	39	22
9	45	25
10	14	8
11	11	8
12	11	5
13	2	1
14	4	3
15	7	7
16 or more	7	6
Total	417	221

This table shows the total number of sample plans based on number of investment choices offered (excluding company stock, money market funds, GICs and stable value funds), along with the number of plans within each total that span eight indexes. A plan that spans offers a sufficient set of choices, so that the plan's optimum tangent portfolio is not improved by including one or more of the eight indexes. The eight indexes consist of four domestic equity indexes, one international equity index, two domestic fixed-income indexes, and one international fixed-income index. The sample period covers the five years from January 1997 through December 2001. Monthly return data for the investment choices in the 401(k) plans were obtained from the CRSP databases.

these plans there are more indexes than fund offerings. However, it is possible that a small set of funds spans the larger set of indexes, either because some of the indexes are not desirable investments or because some of the funds are combinations of two or more of the indexes. However, this does not happen for many plans offering a small set of investment choices. For plans holding seven or more funds, we find that about 60% of the plans offer investment choices that span the relevant space investors are interested in.[14] Of course, the glass is also about half empty in that 40% of the plans may leave investors unsatisfied. Finally, it is not until plans offer 14 or more investment choices (4.3% of all plans) that virtually all plans offer investment choices that span the space investors should be interested in. Of the 417 plans, only 53% span the space obtainable from the eight indexes.[15] While some 401(k) plans offer participants a rich

[14] The sample of 417 plans was constructed to include only those 401(k) plans where all offerings had five years of history. The distribution of the number of offerings with that restriction differs from the distribution of the number of offerings by 401(k) plans in general. If we apply the distribution of those plans that span to the distribution of investment choices shown in Table 1 and assume that all plans with 17 or more investment choices span, the percentage of plans with seven or more offerings that have choices which span falls to 56%. Huberman and Jiang (in press) provide evidence that investors with a wider array of choices only pick three or four funds. These investors may not be taking advantage of the spanning opportunities available for plans with large number of offerings.

[15] For the reasons discussed in the prior footnote, we apply the distribution of plans that span to the distribution of investment choices shown in Table 1, counting each plan offering 17 or more investment choices as a "yes." Applying these rules, the percentage of plans that span is 47%. If plans systematically dropped funds with poor performance, the funds remaining in the plans would have higher returns and more funds would appear to span.

enough selection of investment choices to satisfy their needs, clearly a number of 401(k) plans do not do so.[16]

Before leaving this section, it is worthwhile examining the loss in return to 401(k) plan holders due to plans not spanning the relevant space. To measure this we employ the Sharpe ratio. Recall that the Sharpe ratio is equal to the return on a particular portfolio minus the riskless rate, all divided by the standard deviation of the portfolio. The Sharpe ratio is the standard measure of the efficiency of a portfolio in mean standard deviation space. Two portfolios with the same ratio are equally desirable. Thus for each plan in our sample we can ask the question: given the optimum portfolio for a plan that does not span, how much higher would the return have to be to give the same Sharpe ratio as the optimal portfolio composed of the eight indexes? In order for the 196 plans in our sample which did not span to have the same Sharpe ratio as the optimum portfolio comprised of the eight indexes, the average return on the plans' optimum portfolios, holding risk constant, would have to increase by 0.178% per month. This means that monthly returns would have to increase from 0.866% to 1.044% per month. For an investor who remains in a plan for 20 years, this results in ending wealth that is 53% higher than he or she would receive in a plan with insufficient options.[17] Thus, investors in 401(k) plans are sacrificing significant return because plan administrators are offering an incomplete set of investment alternatives.[18]

We will now examine the characteristics of mutual funds selected by pension plans to see how that impacts plan characteristics.

3. Characteristics of the specific mutual funds selected

In this section we will explore the characteristics of the funds selected by plan administrators given the number and types they select. We will initially explore whether the funds they selected have positive alphas. We will then explore the risk characteristics of the funds selected.

3.1. The risk-adjusted performance of plan funds

The analysis up to this point has been concerned with whether pension plans offer participants adequate types of choices. A second and very interesting question is: given the types of choices offered to participants, is management selecting individual mutual funds that outperform random selection from these types? Even if management is offering the plan participant enough choices, the investor may be forced to choose from among mutual funds that are dominated by other funds of the same types not offered by the plan. To determine this, we need to construct a model to measure performance.

[16] As a further check on plans spanning, we considered whether plans spanned the space of the simplest set of choices we could think of: a broad stock market index (the Wilshire 5000 index), a broad bond market index (a combination of the Lehman U.S. Government/Credit index and the Lehman Mortgage-Backed Securities index), and an international index (the MSCI EAFE index). We adjusted the returns of the 3 indexes to reflect normal management fees (just as we did for the 8 indexes). With this limited set of 3 indexes, more plans offered choices that spanned the indexes' space (236 of the 417 plans).

[17] Ending wealth would be $[1.00866]^{240}$ versus $(1.01044)^{240}$ or 7.92 versus \$12.09 for every dollar invested.

[18] These differences are much larger than any possible differences due to expense ratios between index funds and active portfolios. See Elton, Gruber and Blake (1996) for estimates of expense ratios.

For each mutual fund i ($i=1,\ldots,n$), we will use the alpha (α_i) estimated from a multi-index model to measure the fund's performance. Alpha is the intercept of the following time-series regression:

$$R_{it} - R_{Ft} = \alpha_i + \sum_{k=1}^{K} \beta_{ik} I_{kt} + \varepsilon_{it} \tag{3}$$

where

1. R_{it} is the return on mutual fund i in month t;
2. R_{Ft} is the riskless rate in month t;
3. I_{kt} is either the return on index k in month t if the index is the difference between two return series, or the excess return on index k in month t above R_{Ft} if the index is based on a single return series (the actual indexes are described below);
4. β_{ik} is the sensitivity of fund i to index k;
5. ε_{it} is the random error term for fund i in month t.

Mutual funds were divided into three types: stock funds, bond funds and international funds. For stock funds we used a five-index model: the excess return (over the riskless rate) on the S&P 500 index, the return on the Fama–French small minus big (SMB) factor, the Fama–French high minus low (HML) book to market value factor, the excess return on the Lehman Government/Credit index, and the excess return on MSCI EAFE index. This is similar to the model used by Elton, Gruber and Blake (1999). Two indexes require some comment. First, the bond index is needed both because the stock category includes many funds that are combinations of bonds and stock, such as balanced and income funds, and because funds in common stock categories such as aggressive growth or long-term growth often hold part of their portfolios in long-term bonds. Failure to include a bond index imputes to alpha any return on long bonds different from the riskless rate. The other non-standard index is the international index. During the period of this study many domestic stock funds included international stocks in their portfolios and failure to include an international index could cause alpha to include the effects of a fund's international holdings.

For bond funds we used the excess return on the following four indexes: the Lehman Government/Credit index, the Lehman Fixed Rate Mortgage-Backed Securities index, the Credit Suisse/First Boston High-Yield Index, and the Salomon Brothers Non-Dollar World Government Bond Index. The first three indexes are supported by the work of Blake, Elton and Gruber (1993). The addition of the international index is needed to capture the tendency of some bond funds to include international bonds in their portfolio over this period.

Finally, for international funds we used the excess return on the following five indexes: The S&P 500 index (since world funds invest in part in the U.S.), three MSCI indexes (Europe, Pacific and Emerging Market), and the Salomon Brothers Non-Dollar World Government Bond Index.[19]

We computed alphas for 27 months following the date of our sample. (27 months is the longest period for which we have data.) Using subsequent return data to evaluate performance

[19] The indexes used here are similar to those used earlier except we have separated out the mortgage-backed index and we use a finer breakdown on the international stock index to be consistent with performance measurement practice.

Table 4
Performance of 401(k) plan mutual funds

Type	Number of distinct funds	Average alpha	Average differential alpha	Average differential expense	Expense adjusted differential alpha
Stock funds	326	−0.207%*	0.022%	−0.033%*	−0.011%
Bond funds	90	−0.060%*	0.043%*	−0.028%*	0.015%
International funds	65	−0.083%**	0.064%	−0.038%*	0.027%
Average	481	−0.163%*	0.031%**	−0.032%*	−0.001%

*=significant at the 1% level; **=significant at the 5% level.
This table shows average monthly alphas from multi-index models (Eq. (3) in the text), where the indexes used depend on the type of fund (stock, bond or international). For stock funds, the indexes are the excess return on the S&P 500 index, the Fama–French small minus big (SMB) and high minus low (HML) factors, the excess return on the Lehman Government/Credit Index, and the excess return on the MSCI EAFE index. For bond funds, the indexes are the excess return on the following indexes: Lehman Government/Credit Index, the Lehman Fixed Rate Mortgage-Backed Securities index, the Credit Suisse/First Boston High-Yield index, and the Salomon Brothers Non-Dollar World Government Bond index. The indexes used for international funds are the excess returns on the following indexes: the S&P 500 index, three MSCI indexes (Europe, Pacific, and Emerging Markets) and the Salomon Brothers Non-Dollar World Government Bond index. "Differential alpha" is the alpha on a plan's fund less the average alpha on a set of randomly selected funds of the same type as the plan's fund. "Differential expense" is the monthly expense ratio of a plan's fund less the average monthly expense ratio form a set of randomly selected funds of the same type as the plan's fund. "Expense adjusted differential alpha" is a plan fund's differential alpha plus a plan fund's differential expense. The performance sample period covers 27 months, starting January 2002 and ending March 2004. Monthly return data for the investment choices in the sample 401(k) plans were obtained form the CRSP databases.

eliminates the bias that would result if 401(k) plans added new funds with superior past performance and past data was used to evaluate them.[20]

Alpha from a multi-index model is widely accepted as a measure of both relative and absolute performance. We made one added adjustment to obtain our performance measure. The overwhelming evidence is that alpha is on average negative for mutual funds. To ascertain whether management is doing a good job of selecting funds, we subtract from the alpha for each fund selected by a 401(k) plan the average alpha from a randomly selected sample of funds from the same category (stock, bond or international). We call this difference the "differential" alpha. The population of funds we use as a comparison consists of all mutual funds that exist as of the end of 2001 and have five years of history. These are the same criteria we used when selecting 401(k) plans to include in our total returns sample.

Table 4 shows the average alphas by category for the funds selected by the 401(k) plan administrators as well as the differential alphas. Although 401(k) plan administrators selected funds with poor performance (negative alpha for each category) they selected funds that had smaller negative alphas than random selection. However, only for one category, bond funds, was the average differential alpha significantly different from zero.

What can account for the superior performance of 401(k) pension fund administrators? From Table 4 it is clear that the expense ratios account for almost all of the difference in performance for the three categories of funds plans hold. We don't know the decision-making process of plan administrators. Thus, we cannot tell whether the lower expenses are because of a choice of the

[20] A few of the funds did not have 27 months of data. In these cases we used its last 27 months to estimate the coefficients of the regression. We then took a weighted average of the relevant alpha on the fund prior to its disappearance (overall alpha plus residuals) and the average alpha of the remaining funds in the plan subsequent to disappearance date. This assumes the investors transferred their investment to the remaining funds after a fund disappears.

plan administrators or are a characteristic of the mutual funds that are marketed to 401(k) plans.[21]

In summary, funds selected by plan administrators have slightly higher excess returns than a randomly selected set of funds of the same type, but the majority of the difference is due to plan administrators holding funds with lower expense ratios.

3.2. Risk characteristics

In this section we will examine the risk of funds held by 401(k) plans compared to the risk of "synthetic" 401(k) plans constructed by using random selection of publicly available mutual funds. Examining the holdings of the actual 401(k) plans shows that almost all plans include at least one bond and one stock fund. Recognizing this we employed the following random selection rule:

1. Require at least one bond and one stock fund.
2. Make the odds of selecting a fund from any ICDI category equal to the proportion that category represents of the holdings of all 401(k) plans.
3. Within a category, make the odds of selecting any fund equal.

To calculate portfolio variances for both the actual 401(k) plans and the synthetic 401(k) plans, we need to formulate a rule to represent the investment weighting for a hypothetical plan participant. Many investors equally weight their 401(k) plan offerings (see Benartzi and Thaler (2001) and Liang and Weisbenner (2002)), so we initially use one divided by the number of a plan's investment choices to represent a participant's chosen investment weight in each of the plan's mutual fund offerings.[22] Calculating the average variance of return for each plan and for random selection under the rules described above, we find that plans have a variance that is 2.08 higher than using random selection and that the difference is statistically significant at the 0.01 level.[23]

Why is risk higher on the portfolio of funds selected by plans? Portfolio risk is affected by the variance of the mutual funds selected and the correlation of returns across funds. The average variance of individual funds selected by plan administrators is lower than the variance of randomly selected funds. However, the correlation is higher. Elton, Gruber and Green (2004)

[21] The average fund held by 401(k) plans is much larger than the mean fund (most are in the top decile by size). Larger funds on average have lower expenses. Differential size accounts for most of the lower expenses.

[22] We exclude from the investment choice sets company stock, GICs, stable value funds and money market funds. Huberman and Jiang (in press) provide evidence that investors in a plan with many choices limit their investment to three or four choices but invest $1/N$ in these choices. The same relationship between the risk on plan funds and on random funds holds when funds offer three or four funds as it does on average. Thus, if investors randomly select three or four funds from a larger set, the conclusions are unchanged.

[23] There are an infinite number of alternative weighting schemes that can be examined, and none have any special empirical support. Most can be constructed as linear computations of the variances calculated using the $1/N$ rule. For example, if the investor's proportions in five funds were 35%, 35%, 10%, 10% and 10%, these proportions can be obtained by combining an equally weighted two-fund portfolio, with 50% in each of the first two funds, with an equally weighted five-fund portfolio with 20% in each of the five funds, where the investment is 50% in each portfolio. The variance of this combined portfolio depends on the variance of the equally weighted two- and five-fund portfolios and the correlation between them. The variances of equally weighted portfolios are generally smaller than those for actual plan portfolios. Further, we know correlations are smaller for random portfolios than for actual plan portfolios. Thus the variance of unequal-weighted plan portfolios should be higher.

have shown that funds within a fund family have higher correlation than similar types of funds across families. In our sample 195 of the 417 plans only offer funds from one family, and 53 offer two families but have only one fund from the second family. It is the concentration of offerings within a family that causes the higher correlation and higher risk.

4. Company stock

The analysis to this point has ignored company stock as an asset in 401(k) plan offerings. We know that there is a tendency for plan participants to place a disproportionate fraction of their plan assets in company stock and that this is harmful. In this section we explore whether including a firm's own stock as one of the investment choices in the 401(k) plan is harmful per se rather than through irrational decisions on the part of participants. This could occur, for example, if plans that included company stock offered inferior alternatives to encourage employees to invest in the company stock.

On average, companies offering company stock as an investment choice offer the same number of mutual fund choices as those that do not offer stock; therefore, companies offering company stock do not offer plan participants fewer fund choices as a mechanism to encourage participants to hold more company stock. In addition, slightly more plans that offer company stock as a choice span.

What does the inclusion of company stock do to the optimum portfolio? Examining spanning tests for the 55 plans which included company stock as an investment choice shows that including company stock as a choice in the plan increases the number of plans that span from 33 to 37. While there is a small increase in spanning, the major question is: how does including company stock affect the desirability of optimal portfolios to the participant? To answer this question, we examined the Sharpe ratios for optimal portfolios for plans offering company stock. When company stock did not enter the optimal portfolio, the average Sharpe ratio was 0.240. When company stock entered the optimum portfolio for plans that offer company stock, the average Sharpe ratio was 0.255.[24] If we compare the increase in Sharpe ratios from including company stock with the increase from including a randomly selected mutual fund instead, the difference is close to zero and is neither statistically significant nor economically significant.

Considering the 401(k) plan as the participant's sole financial asset, the inclusion of company stock in a plan seems to neither improve nor harm the investor making intelligent 401(k) plan choices. However, since a plan participant's labor income may be highly correlated with the performance of the company stock, a portfolio including labor income, 401(k) mutual funds and the company stock may be significantly more risky than a portfolio excluding the company stock. In addition, since research has shown that many individuals hold too large a fraction of their plans' assets in the stock of their own company, either because they chose to do so or because the company forces this position through its matching policy, the inclusion of company stock can unduly increase the risk of a participant's portfolio.

5. Plan characteristics

In this section of the paper we examine the relationship between plan characteristics and performance. Before we turn to performance per se, we want to examine one characteristic of

[24] This is true despite an increase in risk that results from holding company stock. For further discussion of the impact of holding company stock, see Poterba (2003) and Brown et al. (2006).

Table 5
Plan characteristics grouped by plan asset size deciles

Plan asset size decile	Avg. plan asset size (thousands)	Avg. number of investment choices	Percentage of plans that use sophisticated strategies	Percentage of plans offering company stock as investment choice	Percentage of plans that vote proxy
1	$2124.579	4.42	13.16%	7.89%	13.16%
2	$6045.974	5.21	7.69%	7.69%	7.69%
3	$10,857.308	6.21	10.26%	5.13%	15.38%
4	$16,882.821	5.23	7.69%	20.51%	2.56%
5	$24,754.590	6.08	12.82%	20.51%	7.69%
6	$37,363.923	7.28	10.26%	17.95%	12.82%
7	$57,851.154	6.72	17.95%	12.82%	7.69%
8	$88,923.718	7.82	20.51%	38.46%	23.08%
9	$173,890.667	7.85	28.21%	41.03%	15.38%
10	$780,277.821	8.18	46.15%	46.15%	17.95%
Spearman rank corr.	1.00*	0.95*	0.77*	0.84*	0.49

This table shows averages and percentages of plan characteristics when the sample plans are grouped into size deciles using total net assets. The averages shown are based on the number of investment choices excluding company stock, money market funds, GICs and stable value funds. The percentages shown are based on number of plans in a given size decile. The Spearman rank correlations shown are of the decile column with the given column. * = significant at the 1% level. This table is based on 2001 survey data of 417 401(k) plans obtained from Moody's Investor Services.

plans that seems to have a major impact on how management behaves and which serves as a parameter that might affect performance.

In Table 5 we divide all plans by the size of assets invested in each plan into 10 deciles.[25] The average size of the plan in each decile is shown in the second column. There is a wide variation in plan size, with the average plan in the tenth decile over 300 times as large as the average plan in the first decile. The first question we examine is whether plans with more assets under management offer participants more investment choices. As shown in Table 5, there is a clear and statistically significant relationship (at the 1% level) between plan size and the number of investment choices offered. Since from our spanning tests we know that more investment choices are generally better for investors, this suggests that large plans ceteris paribus offer an advantage to the 401(k) participants.

Are companies with large 401(k) plans more likely to employ sophisticated strategies such as utilizing security lending, futures and options, hedging strategies and quantitative methods? As shown in Table 5, a higher percentage of larger plans engage in more sophisticated investment strategies. The relationship with size is statistically significant at the 1% level. We next examine the relationship between size and whether a plan votes proxies in the companies it owns. Proxy voting can be interpreted as either another measure of sophistication or as a measure of social consciousness. We find at best a weak positive relationship, one that is not statistically significant. Finally, we examine the relationship between the size of plan assets and the probability of a company including its own stock in its 401(k) plan. Not surprisingly, large plans show a stronger tendency to include company stock in the plan than do small plans, and this relationship is significant at the 1% level.

[25] We were unable to obtain plan size data for 28 of the 417 tracked plans; the size deciles were formed using the remaining 389 plans.

We now turn to an examination of whether the use of sophisticated strategies improves the position of plan participants. To do so we examine their impact on spanning and average risk-adjusted return. From portfolio theory we know that the greater the number of investment choices offered ceteris paribus, the more likely the offerings will span the space. Thus, to examine spanning question we need to control for number of investment choices. We divided the plans into two groups based on whether they used sophisticated strategies. Within each group, given the number of investment choices offered, we compute differences in actual proportions that span and expected proportions that span. We then compared these differences between the group that employed sophisticated strategies and the group that did not. The difference, while in the expected direction, was not statistically significant. Finally, we examine average differential risk-adjusted returns between plans that employ sophisticated strategies and those that do not. Differential risk-adjusted returns are the risk-adjusted return of the plan fund minus the average risk-adjusted return for the population of funds within the same ICDI category. Risk-adjusted returns are then averaged by plan and then across plans. Those plans that use sophisticated strategies have higher differential alpha than those that do not at the 1% significance level.

6. Conclusion

In this paper we examine the reasonableness of the investment choices offered by 401(k) plans. This is an important subject. The payoff of a pension plan to any investor is the product of two different decisions: what the investor is offered, and what he or she chooses from what is being offered. While a lot of attention has been paid to participant choice, little attention has been paid to the relevancy of the choices offered to participants. If investors are given an inferior set of choices in their plan, the effectiveness of their choice is severely constrained.

The major findings of this paper concern the adequacy of plan offerings. We use spanning tests to see if the plan offerings span the space offered by eight indexes. Only 53% of 417 plans span the space defined by eight indexes. This means that, for 47% of the plans, the plan participants would be better off with additional investment choices. In fact, if these plans spanned the 8 indexes, participants' average return would improve by 2.3% per year, which is 21.8% of the return on an 8-index portfolio with the same level of risk. While significant on a 1-year basis, over a 20-year period (a reasonable investment horizon for a plan participant), the cost of not offering sufficient choices makes a difference in terminal wealth of over 53%. Since, for more than one half of plan participants, a 401(k) plan represents the participant's sole financial asset, the consequences are serious.

We next examine the return and risk characteristics of the mutual funds held by plans. We find that the plans select funds that show negative performance, but the performance is still better than the average performance of similar randomly selected funds. The differential performance between funds selected by plans and randomly selected funds is roughly equal to the difference in expenses.

We next examine risk. If a plan sponsor used a common-sense rule of insisting that the plan include at least one stock and one bond fund, then plan risk from random selection would be smaller than the actual risk of the 401(k) plans. Although the individual funds selected by 401(k) plans have lower variance of monthly return than randomly selected funds, the correlation between them is higher.

We then examine plan characteristics to see what the characteristics of superior plans are. There is a strong correlation between the number of investment choices a plan offers and size. This is a strong indication that participants in larger plans are better off than participants in smaller plans. In addition, larger plans are more likely to use sophisticated strategies in the plan. This raises the question of whether the use of sophisticated strategies improves results for investors. We find that, controlling for plan size, the use of sophisticated investment strategies increases the probability of spanning and results in higher differential return. However, only the difference in differential return is statistically significant.

Finally, we examine the effect of offering company stock as an investment choice. We find that plans that offer company stock on average provide the same number of mutual fund choices as plans that do not offer company stock. The overall evidence is that including company stock does not have a major positive or negative effect on the desirability of the investment choices offered to participants, though as pointed out it can lead to their making inferior investment decisions.

Acknowledgements

We would like to thank Kobi Boudoukh, Joshua Rauh, Robert Whitelaw and seminar participants at the 2004 European Finance Association meetings (Maastricht) and the University of Michigan for helpful comments and suggestions.

References

Agnew, Julie, Balduzzi, Pierluigi, 2003. What do we do with our pension money? Recent evidence from 401(k) plans. Unpublished manuscript, Boston College.

Ameriks, John, Zeldes, Stephen P., 2004. How do household portfolio shares vary with age? Unpublished manuscript, Columbia University.

Angus, John, Brown, William, Smith, Janet Smith, Richard, 2005. What's in your 403(b)? Academic retirement plans and the cost of underdiversification. Unpublished manuscript. Claremont University.

Bekaert, Geert, Urias, Michael S., 1996. Diversification, integration, and emerging market closed-end funds. Journal of Finance 51, 835–870.

Benartzi, Shlomo, Thaler, Richard, 2001. Naïve diversification strategies in retirement saving plans. American Economic Review 91 (1), 78–98.

Blake, Christopher R., Elton, Edwin J., Gruber, Martin J., 1993. The performance of bond mutual funds. Journal of Business 66, 371–403.

Brown, Jeffrey, Liang, Nellie, Weisbenner, Scott, 2006. 401(k) matching contributions in company stock: costs and benefits for firms and workers. Journal of Public Economics 90, 1315–1346.

Chen, Zhiwu, Knez, Peter, 1996. Portfolio performance measures: theory and application. Review of Financial Studies 9, 511–556.

Choi, James, David, Laibson, Madrian, Brigitte, Metrick, Andrew, 2004. Active Decisions: a Natural Experiment in Savings. Harvard University mimeo.

Choi, James, Laibson, David, Madrian, Brigitte, Metrick, Andrew, in press. Defined contribution pensions: plan rules participant choices and the path of least resistance. In McCaffrey and Slemrod. Behavioral Public Finance.

De Roon, Frans, Nijman, Theo E., Werker, Bas J.M., 2001. Testing for mean variance spanning with short sales constraints and transaction costs: the case of emerging markets. Journal of Finance 56 (2), 721–742.

Elton, Edwin J., Gruber, Martin J., Blake, Christopher R., 1996. The persistence of risk-adjusted mutual fund performance. Journal of Business 69 (2), 133–157.

Elton, Edwin J., Gruber, Martin J., Blake, Christopher R., 1999. Common factors in fund returns. European Financial Review 3 (1), 1–23.

Elton, Edwin J., Gruber, Martin J., Clifton Green, T., 2004. The impact of mutual fund family membership on investor risk. Unpublished manuscript, New York University.

Fama, Eugene, French, Ken, 1995. Size and book-to-market factors in earnings and returns. Journal of Finance 50, 131–156.

Glen, Jack, Jorion, Philippe, 1993. Currency hedging for international portfolios. Journal of Finance 48, 1865–1886.

Hansen, Lars, Heaton, John, Luttmer, Erzo, 1995. Econometric evaluation of asset pricing models. Review of Financial Studies 8, 237–274.

Huberman, Gur Wei Jiang, in press. Offering versus choice in 401(k) plans: equity exposure and number of funds. Journal of Finance.

Huberman, Gur, Kandel, Shmuel, 1987. Mean-variance spanning. Journal of Finance 42, 873–888.

Huberman, Gur, Sengmuller, Paul, 2004. Performance and employer stock in 401(k) plans. Review of Finance 8, 403–443.

Investment Company Institute, 2000. 401(k) plan participants, characteristics contributions, and accounts activity (Spring 2000).

Liang, Nellie, Weisbenner, Scott, 2002. Investor behavior and the purchase of company stock in 401(k) plan design. Unpublished manuscript, University of Illinois.

Madrian, Brigitte, Shea, Dennis F., 2001. The power of suggestion: inertia in 401(k) participation and savings behavior. Quarterly Journal of Economics 116, 1149–1187.

Poterba, James, 2003. Employer stock and 401(k) plans. American Economic Review 93, 398–404.

36

Institutional Investors and Mutual Fund Governance: Evidence from Retail – Institutional Fund Twins

Richard B. Evans and Rüdiger Fahlenbrach*

March 2012

Abstract

Some investment advisors offer multiple versions of a fund with the same manager and highly correlated returns. But these "twin" funds are separate portfolios for different investors with differing abilities to select and monitor managers. We investigate whether retail investors benefit from investing alongside their institutional counterparts. We find that retail funds with an institutional twin outperform by 1.5% risk-adjusted annually. We demonstrate that institutional twin investors are more sensitive to high fees and poor risk-adjusted performance than retail investors. Using a matched sample of retail and institutional twin funds, we analyze whether the difference in sensitivities can help explain the better performance by focusing on changes to fees and portfolio composition of retail funds after the creation of an institutional twin. We find that after the institutional twin is created, expenses decrease and measures of managerial effort increase, consistent with the reduction of agency problems from greater monitoring.

JEL Classification: G23, G34

Keywords: Governance; Mutual funds; Institutional investors; Performance sensitivity; Identification

* Evans (evansr@darden.virginia.edu) is Assistant Professor of Finance at the Darden School of Business, University of Virginia. Fahlenbrach (ruediger.fahlenbrach@epfl.ch) is Associate Professor at Ecole Polytechnique Fédérale de Lausanne, and affiliated with the Swiss Finance Institute. We gratefully acknowledge helpful comments and suggestions from an anonymous referee, Amit Goyal, Erwan Morellec, David Musto, Brian Reid, Laura Starks (the editor), Russ Wermers, and seminar and conference participants at Darden School of Business, Drexel University, Escuela Superior de Administración y Dirección de Empresas, Ecole Polytechnique Fédérale de Lausanne, McIntire School of Commerce, Wharton's 2011 Rodney L. White Center Conference on Household Portfolio Choice and Investment Decisions, Universitat Pompeu Fabra, Drexel University, University of Kentucky, and William & Mary. Viktoras Vatinas and Giang Ngoc Hoang provided excellent research assistance. Evans and Fahlenbrach gratefully acknowledge financial support from INQUIRE Europe and Fahlenbrach acknowledges financial support from the Swiss Finance Institute and the National Centre of Competence in Research "Financial Valuation and Risk Management" (NCCR FINRISK). Address correspondence to: Richard Evans, University of Virginia, Darden School of Business, 292A Faculty Office Building, Charlottesville, VA 22906.

Introduction

The ability of investors to vote with their feet is the principal investor safeguard in mutual funds. Because mutual fund investors can redeem their shares at net asset value, they can effectively remove the manager from the control of those assets. Fama and Jensen (1983) liken the feature of 'redeemable claims' to a 'partial takeover or liquidation', and argue that this market governance reduces the need for other forms of governance in mutual funds.

Whether or not 'redeemable claims' effectively safeguard investors, however, depends on whether investors use the correct criteria to evaluate funds, and the existing evidence suggests that retail investors fail to respond to many useful signals.[1] In stark contrast with the evidence for retail investors, sophisticated institutional investors respond to useful measures such as expenses and risk-adjusted performance. They exercise market governance and punish poorly performing managers by withdrawing assets under management (e.g., Del Guercio and Tkac (2002); and Goyal and Wahal (2008)).

In this paper, we examine whether retail investors can benefit from the ability and willingness of institutional investors to exercise market governance. We first show that fund flows of sophisticated institutional investors in our sample indeed respond to useful measures such as expenses and risk-adjusted performance, while retail fund flows are much less responsive. We then examine the performance of a subset of retail mutual funds that offer a separate version of the fund for institutional investors either in mutual fund or separate account form, but where the same managers follow virtually the same strategy for both the retail and institutional assets. We find that retail mutual funds with institutional twins perform better than retail funds without, even after accounting for the endogeneity of the decision to offer an institutional twin. The risk-adjusted excess performance is an economically and statistically significant 1.5% annually.

[1] For example, Sirri and Tufano (1998) and Del Guercio and Tkac (2002) find that mutual fund investors use raw return performance to evaluate funds and flock disproportionately to recent winners but do not withdraw assets from recent losers. This convexity leads to well-known problems such as mutual fund managers having incentives to alter the risk of their portfolios if they are close to being among the winners (e.g., Brown, Harlow, and Starks (1996); and Chevalier and Ellison (1997)).

In additional tests, we use cross-sectional differences in the creation dates of the institutional and retail mutual funds to examine whether performance for retail mutual fund investors improves after the addition of institutional assets. Specifically, we focus on the quarter of our twin matches (28.5% or 132 out of 463) where the institutional twin is created after the retail fund. To address this issue, we use propensity-score matching techniques to compare the change in performance for retail twin funds before and after the creation of their institutional twin (treatment group) with a carefully matched sample of funds with no institutional twin (control group). The risk-adjusted performance of the retail funds improves economically and statistically significantly after the addition of the institutional twin relative to a propensity-score matched control sample. In this matched sample, the relative performance improvement of retail funds with twins is driven by a deterioration of performance of the control sample, while the retail funds with twins have stable performance. The stable performance is surprising, because the institutional assets increase total assets under management in the investment strategy (i.e. retail and institutional assets combined) by over 200 percent, and the past literature typically finds that performance is a decreasing function of fund size (e.g., Chen, Hong, Huang, and Kubik (2004)).

We then examine different channels through which monitoring by investors in the institutional twin fund could improve the performance of the retail twin funds. For reasons of both convenience and legal liability[2], retail and institutional versions of the same fund hold virtually the same portfolio and the level of expenses in one portfolio can affect levels in the other. As a result, investors in the retail twin could potentially benefit from increased monitoring by institutional investors if that monitoring resulted in reduced fees or increased managerial effort. To test this, we use the propensity matched sample design described above to examine the change before and after the creation of an institutional twin for three channels: direct expenses, indirect expenses and manager effort. To examine the first channel, direct expenses, we look at changes in the expense ratio and we find a small but statistically significant decrease in the treatment group expense ratio relative to the control. For the second channel, indirect expenses (i.e.

[2] As we discuss in more detail in Section 4, the case of Young vs. Nationwide Life Insurance established the basis of fund liability for differences in performance between twin funds.

those costs that are not included in the expense ratio but are subtracted from fund assets including brokerage commissions, trading/implementation costs, etc.), we examine the change in the return gap measure of Kacperczyk, Sialm and Zheng (2008). The return gap is the difference between a fund's actual return and the return on a hypothetical buy and hold portfolio of the fund's most recently disclosed holdings net of expenses. As such, the return gap abstracts from direct costs, but measures indirect costs such as brokerage commissions, trading costs and other implementation costs. Using this measure, we find that the return gap for retail funds with institutional twins improves strongly relative to the control sample, accounting for approximately 1/3 of the overall increase in risk-adjusted performance. We are also able to measure two specific dimensions of these indirect costs: soft-dollar usage and brokerage commissions. We find that the use of soft dollar payments for distribution decreases after a twin fund was created and we obtain qualitatively similar results when we analyze the brokerage commission rate. One important caveat of the return gap analysis is that in addition to measuring indirect costs, return gap is also a short-horizon measure of manager effort, capturing short term performance enhancement/degradation due to trading. Consequently, the return gap result is also consistent with increased manager effort translating to improved fund performance.

While the third channel, manager effort, is difficult to measure directly, we examine the active share measure of Cremers and Petajisto (2009) and Petajisto (2010).[3] Active share measures the differences of a fund's holdings compared to the holdings of its closest benchmark, with more active funds having a higher active share. We find that after a retail fund adds an institutional twin, its active share increases relative to a control group. In addition, managers of these funds select stocks with lower analyst coverage relative to the control group, which is also consistent with higher managerial effort.

Over the past few years, the role of fund governance has attracted considerable attention in both the finance and legal literature. The focus of the empirical finance literature on fund governance has been on internal governance mechanisms such as director equity incentives or the quality of the board of

[3] Active Share represents the share of portfolio holdings that differ from the benchmark index holdings. Cremers and Petajisto (2009) show that funds with high Active Share outperform those with low Active Share.

directors.[4] Our study contributes to this strand of the literature, because the existence of an institutional twin fund can be interpreted as an external governance mechanism. The institutional twin funds consist of significant investments by a small group of institutional investors, comparable to large shareholders in public corporations. To the best of our knowledge, the issue of the importance of large shareholders has not been addressed in the context of mutual funds. There is ample empirical evidence, however, that institutional investors in large public corporations influence firms' corporate governance. The influence can come either indirectly via "voting (or threatening to vote) with one's feet", as the act or threat of selling shares can have disciplinary effects on companies that lead to changes in governance (e.g., Parrino, Sias, and Starks (2003)) or directly via active monitoring (e.g., adoption of anti-takeover provisions (Agrawal and Mandelker (1990)), CEO compensation (Hartzell and Starks (2003), and Almazan, Hartzell, and Starks (2005)), or proxy voting (Gillan and Starks (2000), and Del Guercio, Seery, and Woidtke (2008)).

The issue of market governance as it relates to the so-called "Gartenberg" standard has been an issue of interest in the legal literature as well.[5] The U.S. Court of Appeals, Seventh Circuit and the U.S. Supreme Court discussed the issue of twin funds in their consideration of a recent mutual fund fee case, Jones v. Harris Associates. The case addresses a twin-fund arrangement where Oakmark (the fund family advised by Harris Associates L.P.) charged different fees to retail clients as compared to institutional

[4] Chen, Goldstein, and Jiang (2008), Cremers, Driessen, Maenhout and Weinbaum (2009), and Meschke (2008) all document that funds whose directors hold a larger fraction of their shares exhibit superior performance. Research on mutual fund board governance has identified an impact of boards on fees, stale pricing, fund merger activity, manager turnover and fund performance. Tufano and Sevick (1997) find evidence that fees are lower for mutual funds whose boards are smaller and have a larger fraction of independent directors. Zitzewitz (2003) shows that the incidence of stale-pricing in fund complexes is higher for funds with fewer independent directors. Khorana, Tufano, and Wedge (2007) examine a sample of fund mergers and conclude that there is a higher probability of a merger if a fund has underperformed and if it has a higher fraction of independent directors. Ding and Wermers (2005) find that funds with a larger number of outside directors are more likely to replace a poorly performing manager. Adams, Mansi, and Nishikawa (2010) find that index funds with smaller boards, boards with inside directors who are also fund sponsor officers, and boards made up exclusively of independent directors are associated with improved performance. Dann, Del Guercio, and Partch (2003) find that closed-end investment companies with smaller and more independent boards carry out value-enhancing restructurings.

[5] Excessive fee litigation cases have relied on the Second Circuit's Court of Appeals 1982 opinion in Gartenberg vs. Merrill Lynch Asset Management, Inc. In its opinion, the court established a heavy burden on mutual fund investors. Investors have to prove that the fee charged is "so disproportionately large that it bears no reasonable relationship to the services rendered and could not have been the product of arm's-length bargaining." For an overview of the legal literature on fund governance, see Morley and Quinn (2010).

clients even though both were getting essentially the same investment product. At the core of the argument was how efficient market governance can be for retail mutual fund investors.[6] Our evidence contributes to this literature. It suggests that despite being charged higher fees than the institutional investors, retail twin fund investors can still benefit from a twin fund arrangement *relative* to non-twin retail fund investors.

Our paper is also related to two recent papers on the side-by-side management of hedge funds and mutual funds (Nohel, Wang, and Zheng (2010), and Cici, Gibson, and Moussawi (2010)). The two papers also exploit the fact that some managers manage different funds at the same time, but their focus in on the quality and retention of the manager, and internal transfer payments.[7]

The remainder of our paper is structured as follows. Section 1 describes the data and offers summary statistics. Section 2 examines market governance by comparing the flow-performance and flow-fee sensitivities of institutional fund investors to those of their retail counterparts. Section 3 contains the empirical analysis of fund performance of retail mutual funds with institutional twins. Section 4 examines potential channels through which monitoring by institutional investors could help the performance of twin retail mutual funds. Section 5 examines the robustness of our main results by carrying out a placebo experiment and repeating our analysis with institutional share classes separated out, and Section 6 concludes.

---

[6] On appeal, chief justice Frank Easterbrook not only concurred with the district court's decision for summary judgment, but he challenged the whole premise behind the Gartenberg standard by suggesting that market governance was an effective mechanism to safeguard investors from excessive fees: "The trustees (and in the end investors, who vote with their feet and dollars), rather than a judge or jury, determine how much advisory services are worth." In the dissenting opinion on the petition for rehearing the case En Banc, Judge Richard Posner disagreed with this sentiment stating: "The panel bases its rejection…mainly on an economic analysis that is ripe for reexamination…. Competition in product and capital markets can't be counted on to solve the problem…."

[7] Nohel, Wang, and Zheng (2010) examine 112 cases where the same fund manager simultaneously manages mutual funds and hedge funds. The main finding is that the best mutual fund managers would potentially leave their mutual fund families and open their own hedge funds because of a more attractive compensation package, but that the permission to run an in-house hedge fund works well as a retention device. Cici, Gibson, and Moussawi (2010) study mutual fund performance when parent firms (but not necessarily the same manager) simultaneously manage hedge funds and focus on the inherent conflicts of interest to transfer performance from mutual funds to hedge funds. They find that the mutual funds managed by these firms underperform a matched sample of other mutual funds.

1. Data

Our sample consists of domestic U.S. equity mutual funds[8] in the Morningstar database from January of 1996 to December of 2009.[9] The principal sample used throughout the paper consists of retail funds and our analysis is performed at the fund level (e.g. share classes are aggregated and all variables are value-weighted by the total net assets of the individual share classes of the fund). We classify funds with all retail or both retail and institutional share classes as retail.[10] Table 1 contains sample summary statistics. Panel A of Table 1 contains a breakdown of the number of funds and observations by year. Because our regressions control for lagged variables, the first year of data for our analysis is 1997. The number of funds is 760 in 1997 and increases in almost every year to a maximum of 1,964 in 2008. Overall, our sample contains 2,660 unique retail mutual funds.

Panel B shows the breakdown of funds into those with and without an institutional twin. To construct the sample of possible institutional twins, we combine separate account and institutional mutual fund data from Morningstar. While the separate account data comes directly from Morningstar, to identify the institutional mutual funds we use an internal Morningstar share class identifier and classify a fund as institutional only if all of the fund's assets are institutional investments as designated by Morningstar. We then compare the retail fund sample to the institutional sample to identify twin matches. A twin match is identified if the retail and institutional fund have the same manager(s)[11], investment

[8] We use the Morningstar "U.S. Broad Asset Class", "Global Category" and "Investment Objective" variables to identify the domestic equity sample. Specifically, we require all funds to have a U.S. Broad Asset Class designation of U.S. Stock and an Investment Objective classification of aggressive growth, growth, growth & income, equity income, or small company. We then remove all funds with a Morningstar "Global Category" classification of Real Estate or any domestic equity sector classification. These filters remove fixed-income, real-estate, commodity, international equity, balanced and sector funds from the sample.

[9] The start date of January 1996 coincides with the first date of survivorship-bias free Morningstar data availability.

[10] We categorize mutual funds that have only retail share classes and mutual funds that have both institutional and retail share classes as "retail" funds both to be conservative (classifying funds with some institutional share classes as retail should bias the empirical examination against finding any difference between retail funds with and without twins) and due to the concern that institutional investors in separate accounts and purely institutional funds are fundamentally different from investors in institutional share classes, the latter being less financially sophisticated. In the robustness section, we discuss the impact of relaxing this assumption on our results.

[11] Because fund manager information is used in identifying fund twins, those funds that are missing manager information or only classify the manager as "Management Team" are removed from the sample.

objectives, fund families and a gross return correlation of 0.95 or greater.[12] As Panel B shows, 463 out of 2,660 unique retail mutual funds or 17.4% of the sample observations have an institutional twin. Of these institutional twins, 345 are from the separate account sample and 118 are from the institutional mutual fund sample described above.

Panel C offers summary statistics for the retail funds in the sample, split by whether they ever have an institutional twin or not. The table lists the mean, median, 25th and 75th percentiles of the fund family size (measured as total net assets (TNA) under management), fund size (TNA), expense ratio, turnover (the minimum of fund purchases and sales divided by fund TNA), the quarterly net flow into the fund, the fund's 4-factor alpha (the three Fama and French (1993) factors combined with the momentum factor of Carhart (1997)), and the percent of the observations coming from broker-sold funds. Unless otherwise noted, statistics are based on monthly observations. Fund family size[13], expense ratio, turnover and the percent of observations from broker-sold funds are comparable across the two samples. Funds with twins have larger median TNA, higher flows and better 4-factor alpha performance. While the difference in performance is particularly interesting given the focus of our study, it is important to recognize that this result does not suggest causality. Indeed, we would expect that the funds that are most likely to be sold to multiple clienteles would be those with the best performance. The larger median size of these funds and the higher inflows are also consistent with this interpretation. The summary statistics for the "Twins Sample" also includes information about the start date of the retail fund relative to its institutional twin. The second-to-last row of Panel C of Table 1 gives the mean, median, 25th and 75th percentiles of the relative start date, defined as the difference in years between the inception date of the institutional twin and the inception date of the retail fund. The mean relative start date value of 0.9, for example, indicates that the average institutional twin was started 0.9 years after its retail counterpart and

---

[12] While setting a lower bound on the return correlation is a logical safeguard, when matching on the first three criteria alone, the mean and median return correlation is 0.98 and 0.99.

[13] While the differences in mean and median fund family size between funds with and without an institutional twin are in the billions of dollars, these differences are small relative to the dispersion of fund family size (~$20 or $30 billion).

the negative 25th percentile shows that the sample also consists of twin pairs where the institutional fund was started before its retail twin.

Panel D further explores this heterogeneity in twin start dates, showing the number of twins where the retail fund was started first, the institutional fund was started first, and the retail and institutional funds were started on the same date (Same Incept. Date). In some cases the retail or institutional twin fund was started before the sample period began (1996) so the last three columns of Panel D repeat the breakdown for the sample of 313 twin funds where the inception date for the twin fund (i.e. the second fund created) occurs during the sample period. The 132 twin pairs where the retail fund was started first and the institutional twin was created during the sample period are important for our identification strategy in Sections 3.3 and 4 of the paper. A pre-twin period for the retail fund allows us to identify the incremental effect of the institutional fund's creation on retail mutual fund performance. In addition, such a sample enables us to provide estimates from propensity-score matched samples. With propensity-score matching, we first analyze which retail funds are the most likely to create an institutional twin, and then compare a treatment sample of twin funds with an appropriately matched control sample.

While the majority of our analysis focuses on retail funds, Table 2 characterizes the institutional twin fund sample. Panel A offers summary statistics for the sample of institutional twin fund-months. Because separate accounts do not have the same disclosure requirements as mutual funds, some variables, such as turnover, are not as well populated in the database and the flow and TNA data is only given at a quarterly frequency. As a result, in addition to the mean, median, 25th and 75th percentiles, we also list for each variable the number of fund-month observations available to calculate the statistics. Comparing the institutional funds to their retail counterparts (twins sample in Table 1, Panel C), we see a number of differences. Consistent with an institutional clientele, expense ratios are lower and fund size is larger. Although performance is better for the institutional twins, the difference between the retail and institutional fund performance is roughly equivalent to the difference in expense ratios.

While comparing Panel C of Table 1 and Panel A of Table 2 gives an approximation of the differences between retail funds and their institutional twins, we formally test these differences in Panel B of Table 2. To construct this table, we merge monthly retail fund data with the corresponding monthly institutional twin data when it is available.[14] For each matched pair-month observation, we then calculate the difference in variables of interest. The differences are first averaged over time for each of the matched pairs. Panel B of Table 2 provides cross-sectional sample statistics from the time-series averages of the matched pairs. Columns 5 and 6 of Panel B provide p-values for tests of differences in means (t-test) and medians (sign test), and also shows the number of matched pairs for which the variable of interest is available.

Panel B of Table 2 shows that the average (median) institutional fund is $547.9 ($86.8) million larger than its retail twin and has an expense ratio that is 0.42% (0.44%) lower. Both of these differences are statistically significantly different from zero. In each case, they are also consistent with differences in the type of investors. Because investors in separate accounts typically have much larger investments, they also have much lower expenses. Institutional twins have higher average quarterly net flows. Because the flows from these investors are typically more predictable and less volatile, managers do not need to hold as much cash on hand to meet redemptions and do not need to trade as much to account for them.

Panel B of Table 2 also compares three different annualized performance measures: 4-factor alphas, net total return, and gross total return. The 4-factor alphas are calculated over the previous 36 months with the standard set of factors proposed by Fama and French (1993) and Carhart (1997): market (MKT), market capitalization (SMB = small minus big), book-to-market (HML = high minus low) and momentum. While the institutional funds outperform their retail twins in terms of both 4-factor alpha and total return (both net of fund expenses), comparing the means and medians of these differences to the means and medians of the differences in gross total return, we see that the outperformance is roughly equivalent to the difference in expense ratios (~50 bps). We do find, however, a small but marginally statistically significant difference in mean gross returns, but this is not surprising given the lower turnover and the higher percentage invested in common stock described above.

[14] While return data for the separate accounts is available at a monthly frequency, the total net asset and flow data is only available at a quarterly frequency. For these variables, we calculate the institutional-retail differences at a quarterly frequency.

We finally examine differences in the factor loadings between matched pairs to assess differences in risk. The comparison of the factor loadings on the market, size, and value factor indicate that there are no significant differences across the institutional and retail twin funds.

Overall, the comparisons of Table 2, Panel B suggest that while there are important economic differences between retail funds and their institutional twins, there is little or no difference in the performance other than the fee differential and little or no difference in risk between the twin pairs.

2. Fund-flow-sensitivity of retail mutual funds and institutional mutual funds

In this section, we examine the determinants of net flows into retail and institutional funds to identify whether these two types of investors respond to different signals. Our working hypothesis is that institutional investors use more sophisticated criteria to evaluate fund managers and have a greater aptitude for market governance. We provide two sets of results. Specification 1 of Table 3 is based on the entire sample and thus allows a comparison with earlier papers (e.g., Del Guercio and Tkac (2002)). Specification 2 of Table 3 uses the matched sample of retail and institutional twin funds only. Using a matched pair with the same manager, the same investment objective and the same fund family, we are able to control for the innate ability of the manager and the influence or impact of the fund family on flows.

Specification 1 of Table 3 contains an analysis of the determinants of institutional and retail quarterly flows for all retail and institutional domestic equity funds. The dependent variable is percentage quarterly net fund flow for the next quarter $(t = 0 \text{ to } t = 3)$. The independent variables include the lagged $(t = -1)$ natural log of fund family TNA and fund TNA, the lagged fund expense ratio, lagged turnover, the concurrent $(t = 0 \text{ to } t = 3)$ percentage quarterly flow to different funds with the same investment objective, and the lagged percentage quarterly fund flow. We use 36-month total return and 4-factor alpha computed from the previous 36 months of data as our raw and risk-adjusted performance measure, respectively. We include both the risk-adjusted and total return measure to see whether retail and institutional flows respond more strongly to either measure. Similar to Sirri and Tufano (1998), we allow for non-linearity in the performance measures in all specifications. We use a piece-wise linear performance specification with a kink at the 20th and 80th percentile of returns.[15] The specification has separate retail and institutional coefficients for each variable. The coefficients for retail funds are under the *Retail Coef* header and the coefficients for institutional funds are under the *Inst. Coef* header. We also test whether the coefficients on the expense ratio and performance measures for retail and institutional funds are statistically significantly different from each other. The p-values from these tests are listed at the bottom of the table.

The main focus of our analysis is on the coefficients of the expense ratio and performance metrics. Carhart (1997) provides evidence that past expenses negatively relate to future returns. As a result, investors should avoid high fee funds. We see from specification 1 of Table 3 that institutional investors are almost six times as sensitive to expenses (-0.0079 vs. -0.0468) as retail investors and that the difference is statistically highly significant at a p-value of 0.005.

Carhart (1997) shows that poor risk-adjusted performance is persistent and that funds with high total returns exhibit mean reversion. As a result, investors should avoid funds with poor risk-adjusted performance and they should not chase past total returns. Retail investor flow is convex in both total return and our measure of risk-adjusted performance, 4-factor alpha. These results are consistent with the results of Sirri and Tufano (1998), who show a similar non-linearity with flow responding more positively to high returns (greater than 80th percentile) than to low returns (lower than 20th percentile). In stark contrast, institutional investor flow is unrelated to good raw performance and is more sensitive to poor risk-adjusted performance (4-Fctr Alpha Low) than to good (4-Fctr Alpha High). As the p-values at the bottom of the table indicate, the difference between the retail and institutional response to poor risk-adjusted performance is statistically significantly different from zero. For total return, there is no

[15] For the total return measure, the percentiles are calculated within date and investment objective similar to Sirri and Tufano (1998), while the 4-factor alpha percentiles are calculated within date only. With these percentiles, the formula for the low return is LowRet=Min(0.2, RetPtile); the formula for the medium return is MedRet=Min(0.6, RetPtile-LowRet) , and the formula for the high return is HighRet=RetPtile-MedRet-LowRet.

statistically significant response to good raw performance for institutional fund flow. These results are consistent with the results of Del Guercio and Tkac (2002). The control variables have the expected signs. The coefficient on family size is positive and on fund size negative. This is consistent with larger fund family size proxying for higher visibility or lower search costs for investors. The negative sign on fund size is consistent with Berk and Green's (2004) and Chen, Hong, Huang and Kubik's (2004) diseconomies of scale arguments. The positive and significant coefficients on investment objective flows and lagged fund flows are consistent with previous evidence of herding behavior (i.e., Sirri and Tufano (1998)) within an investment objective and strong positive fund flow autocorrelation generated by automated investment programs such as 401(k), 403(b), 529 or other tax deferred investment programs. We also include the concurrent flow to the fund's twin (Twin Fund Qtrly Pct Flow) in the regression to control for transfers from retail to institutional funds by the same investors or unobserved flow determinants common to both retail and institutional investors. The coefficient on concurrent flow to the fund's twin is positive and significant, suggesting some commonality in both retail and institutional flows that our other independent variables do not capture.

Specification 2 of Table 3 shows the results from regressions examining the determinants of fund flow in the matched sample only. Comparing the expense ratio coefficients for institutional and retail funds, we see that institutional flows are ten times more sensitive to expenses (-0.060 vs. -0.0063) as retail funds and that the difference is statistically significant with a p-value of 0.009.

In the matched sample specification, retail flows are sensitive to both the high and low total return coefficients, and institutional flows are not statistically related to total returns. Looking at the 4-factor alpha coefficients, we again see that the institutional flows are not only more sensitive to poor risk-adjusted performance than retail flows, but also that the piece-wise linear specification for institutional flows exhibits concavity in stark contrast to the observed retail flow performance convexity. In other words, institutional investors respond with greater sensitivity to poor risk-adjusted performance than they do to good past risk-adjusted performance and the difference in the retail and institutional response to poor risk-adjusted performance is strongly significant. The coefficients on the control variables are generally of the same magnitude as in the first specification, with lower statistical significance because of the lower number of observations.

Overall the evidence from flows is compelling. Institutional investors respond more sensitively to variables that predict returns, namely, expenses and poor risk-adjusted performance. In addition, they avoid total return chasing behavior to a greater degree than their retail counterparts. Given this evidence, it is possible that institutional investors play an important role in disciplining fund managers through market governance.

3. Performance results

The previous section demonstrates that institutional fund investors indeed exhibit stronger market governance. But do retail investors benefit from the presence of institutional investors via twin fund arrangements? In this section, we estimate three performance regressions to answer the question. We first provide evidence in Section 3.1 that retail funds with institutional twins have better monthly performance using the entire sample of retail mutual funds, carefully controlling for characteristics known to affect mutual fund returns.

Mutual fund families have discretion over whether to offer a twin institutional and retail fund, and hence our baseline specification is subject to endogeneity concerns. In Section 3.2, we estimate a model that first specifies a selection equation for the probability of a fund advisor offering an institutional twin to an existing retail twin, and then includes a Heckman correction in the monthly performance regression in the second stage.

In Section 3.3, we make use of a particular subset of our sample of twin funds to provide additional support for an increase in performance of retail funds with twin institutional funds. In 132 out of 463 twin observations, the retail mutual fund was created before the institutional twin and the institutional twin was created during our sample period. For this sample we can examine various fund characteristics before and after the institutional twin was created. We use propensity-score matching techniques to compare the change in the 36-month risk-adjusted retail fund performance for retail twin funds around the creation of the institutional twin (treatment group) with changes in performance of a carefully matched sample of funds with similar characteristics, but no institutional twin (control group).

3.1 Baseline performance results

We examine the relationship between the addition of an institutional twin and retail fund performance using the full set of retail funds (i.e. retail funds with and without twins). To do this, we estimate a regression of future monthly fund performance on lagged fund characteristics known to affect expected returns and on several institutional – retail twin indicator variables.

The dependent variable is the fund's 1-month forward-looking 4-factor alpha. This is calculated using the factor loadings estimated over the prior 36 months of data $(t-1 \text{ to } t-36)$. Using these factor loadings and the factor realizations for time *t*, a monthly benchmark return is calculated:

$$r_t^{Benchmark} = \beta_{t-1,t-36}^{MKT} r_t^{MKT} + \beta_{t-1,t-36}^{SMB} r_t^{SMB} + \beta_{t-1,t-36}^{HML} r_t^{HML} + \beta_{t-1,t-36}^{Moment} r_t^{Moment}$$

Taking the difference between the monthly fund return at time *t* and the monthly benchmark return gives the 1-month forward-looking 4-factor alpha.

The independent variables include the natural log of fund TNA and fund family TNA, turnover, the expense ratio, an indicator variable equal to one if the fund is broker sold (indicated by the presence of a front or rear load or a 12b-1 fee), and an indicator variable equal to one if the fund is passively indexed.

Our key independent variables measure the existence of an institutional twin. We split the observations of retail funds with institutional twins into three different categories. The "Inst. 1st" variable is equal to one if the retail fund has an institutional twin and that twin was created before the retail fund. The "Retail 1st - After Inst. Fund Created" indicator variable is equal to 1 if the retail fund has an institutional twin, the twin was created after the retail fund and if the date of the observation is after the institutional fund was created (e.g., the retail fund was created on 1/1/2000, the institutional fund was created on 1/1/2004 and the date of the return observation is 1/1/2005). The "Retail 1st - Before Inst. Fund Created" indicator variable is equal to 1 if the retail fund has an institutional twin, the twin was created after the retail fund and if the date of the observation is on or before the date on which the institutional fund was created (e.g., the retail fund was created on 1/1/2000, the institutional fund was created on 1/1/2004 and the date of the return observation is 1/1/2002).

The panel regressions include fixed effects for date (specifications 1 through 3) and fund family (specifications 2 and 3) as indicated at the bottom of the table. Standard errors are clustered by fund and date. The p-value of a difference in coefficients test comparing the before and after versions of the Retail 1st indicator variable coefficients is shown in the last row of the table and measures whether the performance of the retail fund is statistically significantly improving after the creation of an institutional twin.

Table 4 shows the regression results. Looking at the twin indicator results of specifications 1 and 2, the coefficient on "Inst. 1st" is positive and strongly significant in every regression. The interpretation of this result is that retail funds in the sample that were created as clones of an existing institutional fund outperform other retail funds. Because the dependent variable is 4-factor risk-adjusted monthly performance, the coefficients of 0.1086 and 0.1296 indicate that the retail funds outperformed by between 10.86 and 12.96 basis points per month or between 1.30% and 1.56 % per year. This is consistent with a number of different interpretations including the ability of institutional investors to select superior managers/strategies, the use of superior managers by fund families to manage institutional assets or the additional monitoring by investors in an institutional twin resulting in improved performance.

The "Retail 1st - Before Inst. Fund Created" indicator is positive, but not statistically different from 0. The result indicates that retail funds that will eventually create a twin do not have a statistically significantly higher monthly performance than non-twin funds over the period leading to the creation of the institutional twin, once we control for all characteristics. The coefficient on "Retail 1st - After Inst. Fund Created" is positive, and statistically and economically significant. The coefficients of 0.0978 and 0.1198 indicate that the twin retail funds outperform non-twin funds by between 1.18% and 1.44 % per year after the institutional twin is created.

The comparison between the "Retail 1^{st}- Before Inst. Fund Created" coefficient and the "Retail 1^{st}- After Inst. Fund Created" coefficient shows that retail fund performance improves after the creation of an institutional twin. The coefficient on "Retail 1^{st} - After Inst. Fund Created" is statistically significantly larger than the "Before Inst. Fund Created" coefficient in specifications 1 and 2 as indicated by the p-values comparing the difference in the "Before" and "After" indicator variable coefficients listed at the bottom of the table. This result suggests that once the institutional twin is created, risk-adjusted performance of the twin retail fund significantly improves, despite the negative impact of increased assets under management due to the scale diseconomies documented by Chen et al. (2004) and others.

A potential concern with the performance analysis of specifications 1 and 2 in Table 4 is the possibility of a survivorship or backfill bias affecting the results. Because separate accounts are outside the scope of the Investment Company Act of 1940, they are not subject to the same reporting requirements. As a result, the separate account data in the Morningstar database is provided on a voluntary basis and may be subject to both survivorship and backfill biases. Note that the data used for the performance analysis in Table 4 comes from the retail mutual fund database that does not suffer from these biases. However, if the identification of retail funds with twins is based on the separate account twin data which is subject to biases, it may also affect our results for the twin retail fund indicators.

To alleviate survivorship and backfill concerns, we construct a bias-free sample and repeat the return regressions using this sample.[16] The bias-free sample includes observations where the institutional twin is an institutional mutual fund (these are taken from the regular Morningstar mutual fund database

---

[16] Our separate account database is constructed from annual snapshots from the Morningstar Principia separate account database. The annual snapshots eliminate the possibility of survivorship bias because funds that are removed from the database in later periods are included in earlier snapshots. Because the first snapshot of the annual database covers the period of August 2001 to September 2002, but our retail fund sample starts in 1997, any separate account performance observations from before August of 2001 could potentially have a survivorship bias. We therefore classify any mutual fund observations with a separate account twin before August of 2001 as a potentially biased observation. Separate from the survivorship bias concern is that of backfill bias. If investment advisors only report the returns of their separate accounts that are successful and if they are allowed to backfill the returns of those accounts, the twin funds that are matched with these separate accounts may have significant outperformance due to an upwards bias in the separate account performance. Fortunately, the Morningstar separate account database provides the year and month in which each account was added to the database, which allows us to remove backfilled observations.

and are free of survivorship or backfill bias concerns) or a separate account after August of 2001 and after that separate account started reporting to the Morningstar Separate Account database.

Specification 3 of Table 4 shows results using the bias-free sample. Results are qualitatively and quantitatively similar to what we found before. Note that the p-values of the differences in the before/after tests are smaller (but still significant at the 8% level). The weaker statistical power is to be expected, because the survivorship/backfill bias filter removes approximately half of the twin observations in which the retail twin was created first.

3.2 Heckman correction

Khorana and Servaes (1999) show that the decision to open a new mutual fund is related to a number of variables including fund family size, flows to the family and the investment objective and the superior past performance of other fund offerings by the same family. Because these variables may be related to future fund performance, if the decision to open an institutional twin is related to similar factors, it raises the possibility that our baseline specification is subject to endogeneity concerns. To address these concerns, we use two different methods. First, in this subsection we employ the insights of Heckman (1979) to account for the endogenity of the fund family decision to offer a twin fund. Second, in Section 3.3, we use a propensity-score approach, matching retail funds with an institutional twin with like funds without a twin.

Since we are interested in the performance of retail mutual funds with a twin relative to retail mutual funds without a twin (and we observe the performance of both types of funds), in this subsection we use the treatment effects model of Maddala (1983) that builds on the Heckman self-selection model. The treatment effect regressions are estimated using Heckman's (1979) two-step procedure. We first estimate a probit regression in which the dependent variable is equal to one if the retail fund has an institutional twin in month $t-1$, and zero otherwise. The second-stage regression is the monthly performance regression, and includes a Heckman correction term.[17] Because we are interested in the decision to offer an institutional twin to an existing retail mutual fund, we exclude in the regressions of this subsection all sample observations in which the institutional twin was created first.

Table 5, Panel A contains the results of the first stage selection equation, and Table 5, Panel B contains the results of the second stage performance regressions. Panel A shows that the probability that an institutional twin fund is created in any given month is increasing in the past performance of the retail mutual fund as measured by the 4-factor alpha over the past 36 months, and the size of the retail fund. If the fund is broker-sold, the probability increases as well.

Panel B shows the results of the monthly performance regressions that now includes the Heckman correction term (lambda). The coefficient on the correction term is only significant in one of the three specifications. Accordingly, the indicator variable $Twin_{t-1}$ *(=Yes for Retail 1st - After Inst. Fund Created)* continues to be strongly statistically and economically significant. Depending on the specification, the annualized excess performance of a retail fund after an institutional twin was created relative to a no-twin retail fund is between 1.24% and 1.92%, which is similar to the results reported in Table 4.[18]

3.3. Propensity-score matched sample

While the tests in Sections 3.1 and 3.2 use the entire sample of retail funds, we focus in our last test of fund performance on a subset of the sample. For 132 of our 463 twin funds, the institutional twin is created after the retail fund and the creation date occurs during our sample (1996 to 2009). Using the creation of an institutional twin as identification, we compare the change in performance of a given retail mutual fund before and after its institutional twin is created (treatment group) with the change in

[17] For details, please see Maddala (1983) and Greene (2011). If the selection and treatment equations contain the same explanatory variables, identification only comes from the nonlinearity of the inverse Mills ratio (the Heckman correction term). We include the following variables in the first-stage probit model that we do not include in the second-stage regression to help identification: past performance, past inflows, tracking error, and family institutional TNA indicator variable.

[18] We also carry out the analysis separately for the two different twin types: institutional mutual fund and separate account. The performance difference between the after and before retail twin creation indicator is a strongly statistically significant 1.06% in the case of separate accounts and weakly significant 0.88% in the institutional mutual fund case.

performance of an otherwise similar retail mutual fund that does not have a twin (control group). To identify the control sample, we use a propensity score matching technique that was pioneered by Rosenbaum and Rubin (1983) and has been used recently in the finance literature by, for example, Drucker and Puri (2005) and Aggarwal, Erel, Stulz, and Williamson (2009). As we discussed in Section 3.1, if the decision to open a twin fund is related to past flows, the size of the fund family, or other fund characteristics that potentially play a role for the future performance of the fund, our performance result is subject to an important endogeneity concern. The propensity score approach allows us to address this endogeneity concern by modeling the decision to open a twin fund and to select a control group of funds that most closely resemble our treatment group on these dimensions.

In the first stage of the analysis, we calculate each firm's propensity score, which is equal to the probability that a retail mutual fund with given characteristics creates a twin institutional fund in the next year. In the second stage of the analysis, each retail mutual fund that chooses to create an institutional twin fund (the treated group) is matched with retail mutual funds that have the closest propensity scores, but did not choose to create an institutional twin (the control group).[19] We then calculate the average change in the 36-month 4-factor alpha across the event for the treated group and compare it with the average change in the 36-month 4-factor alpha of the control group.

The coefficients for the first-stage probit regression are given in Panel A of Table 6. The control variables include an intercept, fund performance (4-factor alpha), fund flow, fund size and the size of the fund family's retail and institutional assets under management, an indicator variable of whether or not the family manages any other institutional assets, expense ratio, turnover, tracking error of the fund's returns relative to the 4-factor model determined benchmark and indicator variables for the distribution channel (Broker-Sold) and whether or not the fund is an index fund. The decision to create an institutional twin is related to three factors. Larger past fund flows and a larger fund size increase the probability of creating an institutional twin next year. If a fund is broker-sold, the probability of creating a twin increases as well.

[19] Propensity scores are calculated for all fund-month observations using the coefficients from the propensity score probit model and the control group is chosen from the set of funds with the closest matching propensity score to the treatment group from the same year and month.

It is surprising that the fund families' prior experience managing institutional assets is not significantly related to the probability of offering an institutional twin. It is important to note that the probit regression is *not* estimating the probability of the fund family opening any type of institutional product but rather it is estimating the probability of the fund family opening only a specific type of institutional product: the twin of an existing retail fund. The fund family's institutional TNA variables measure the aggregate assets of all institutional products offered by the family as captured by the separate account and institutional mutual fund databases described earlier and not just twin funds. The majority of institutional assets in the Morningstar database are invested in funds that are not twins. Given the results of Del Guercio and Tkac (2002) and Goyal and Wahal (2008) on the role of performance as a determinant of institutional fund flows and the hiring of institutional managers, respectively, it is also surprising that the risk-adjusted past performance of the fund is unrelated to the decision to create a twin. However, both fund flows and fund size, alternative measures of market demand, are positively related to the probability of creating a twin fund. Additionally, a comparison of the average risk-adjusted performance in the summary statistics for the treatment and control groups in Panel B of Table 6 with the average risk-adjusted performance of the overall population (cf. Table 1, Panel C) shows that the risk-adjusted performance of both treatment and control group is significantly higher.

In spite of the low R-squared of the propensity score model, Panel B of Table 6 shows that the matching works well for all of the fund and family characteristics examined. The differences in past performance, fund flow, expenses, fund and family size and other characteristics between the treatment and control sample are economically small and statistically insignificant.

Table 7 shows the levels of the 4-factor risk-adjusted performance of retail funds in the three years before an institutional twin is created and the three years after an institutional twin is created, as well as the changes in performance around the event. It then compares the levels of and changes in risk-adjusted performance of this treatment group with the levels of and changes in risk-adjusted performance of the propensity-score matched control group. The requirement of a 6-year window decreases the sample size from the original 132 to 98 retail-first funds. For the three years before the institutional twin is created, the average retail mutual fund in the treatment sample has a negative annualized 4-factor alpha of -0.373%. The control group has a risk adjusted performance of -0.170% over the same period and the two are not statistically different (since we matched on past performance in the first stage, this result is not surprising). For the three years after the institutional twin is created, the treatment sample averages risk-adjusted performance of a statistically insignificant -0.038%, and the change in performance across the event is a statistically insignificant 0.335%. The control sample, on the other hand, has a statistically significant negative alpha of -1.383% during the three years after the twin is created, and these funds average a 1.213% deterioration in performance. Comparing the increase in performance for the treatment group of +0.335% with the decrease in performance for the control groups of -1.213%, the retail funds outperform the matched sample after the addition of their institutional twins by a statistically and economically significant risk-adjusted 1.548% per year.

While the econometric methods are very different across sections 3.1 to 3.3, the results from the propensity-score matched sample are very similar to the results on performance improvements we observed in our tests that use the whole sample (1.18% to 1.92%, depending on the specification).

The results we reported in the first set of results are based on nearest neighbor matching using ten observations from the control-group and excluding observations without common support. To show the robustness of our results, we repeat the analysis with different specifications. In the second (third) row, we use a tolerance level on the maximum propensity score distance or caliper length of 0.005 to avoid the risk of bad matches, and then match each treated fund with up to the closest 10 (5) nearest observations. In the fourth and fifth rows, we further reduce the maximum propensity score distance to 0.001. As can be seen from these alternative specifications, our propensity-score matched return results are very robust.

The last block of results repeats the propensity score matching, but excludes all retail twin funds that are subject to potential survivorship and backfill bias. This leaves 53 twin institutional fund creations. Using this sample, we find that the retail mutual funds of the treatment group actually improve performance across the addition of an institutional twin, by a statistically significant 0.54%. Because the performance of the control group deteriorates at the same time, the treatment group outperforms the control group by a statistically and economically significant 1.95%..

The deterioration in performance of the control group in the after period relative to before is not surprising given the evidence in Carhart (1997) that above average fund performance reverts to the mean over time. Sirri and Tufano (1998) and Chen, Hong, Huang and Kubik (2004) point to a likely driver of this performance decrease. Given the superior performance of both the treatment and control group funds (-0.37% and -0.17%) in the 3-year 'before' period relative to the retail fund population average (-1.46%), we would expect these funds to have substantial inflows and the corresponding increase in fund size[20] to negatively affect performance. While the control fund exhibits such a pattern after an institutional twin is created, the treatment group performance does not deteriorate, and in some specifications, there is evidence of a statistically significant improvement in performance. In the next section, we explore possible channels through which institutional monitoring could benefit the performance of the retail fund.

4. Potential channels for increased monitoring

While the evidence in Section 3 about better performance of retail mutual funds with twins is compelling, it does not shed light on the mechanism through which the creation of an institutional twin could benefit retail investors. Using the same propensity score techniques as those used in Section 3.3, we now examine potential channels through which increased monitoring of institutional assets by institutional investors could affect retail twins.

Because the retail fund and the institutional twin are separate portfolios of different size and somewhat different portfolio composition there is no mechanical reason why changes made to the institutional fund's fee structure and portfolio composition would result in changes to the retail fund's fee structure and composition. Legal and regulatory guidelines, however, help us identify why retail investors might benefit from better oversight of the institutional investors.

[20] Comparing the three year average TNA during the before and after periods, both the treatment and control groups have fund size increases of 30 to 40%.

First, the relevant legal precedent for performance differences between twin funds is Young vs. Nationwide Life Insurance.[21] In this case, the shareholders of a variable annuity life insurance fund successfully sued the life insurance fund sponsor on the basis of differences in performance between the mutual fund and its variable annuity fund twin. The case establishes fund liability for differences in performance.

Second, in negotiating fees with the advisor, mutual fund boards can use a comparison of fees charged by other funds or fees charged to other clients such as pension funds or other institutional investors. While the SEC has required boards since 1994 to disclose the material factors used and the rationale for approving an advisory contract (such as a fee comparison)[22], the SEC modified the disclosure requirements in 2004 to specifically require that boards discuss their use of fee and service comparisons, in addition to a small list of other factors. Hence, by allowing lower fees (either direct or indirect) for the institutional fund and not discussing such lower fees when setting the fees of the retail twin fund, a board of directors may expose itself to legal liability and violate its fiduciary duties. However, such a discussion of low institutional fees in the context of retail fees may lead to downward pressure on retail twin fund fees.

Third, in their analysis of the Jones v. Harris Associates case, the Supreme Court makes clear in their reinstatement of the Gartenberg standard that fee comparisons are relevant for excessive fee determinations:

> "First, since the Act requires consideration of all relevant factors, §80a–35(b)(2), courts must give comparisons between the fees an investment adviser charges a captive mutual fund and the fees it charges its independent clients the weight they merit in light of the similarities and differences between the services the clients in question require."[23]

While the focus on much of the legal literature surrounding the Gartenberg standard is on the components of the expense ratio such as advisory fees, the statute is broader and includes any "payments of a material

[21] Young v. Nationwide Life Ins. Co. - 2 F.Supp.2d 914 (S.D. Tex. 1998).
[22] 59 Federal Register 52689 (Oct. 19, 1994).
[23] Jones v. Harris Associates L.P. 559 U.S. ___ (2010)

nature" suggesting inclusion of indirect expenses as well. The court does caution to be "wary of in-apt comparisons based on significant differences between those services", but the nearly identical nature of twin funds suggests that advisors would be particularly mindful of direct and indirect fee differences between twin financial products.

In addition to establishing the legal and regulatory ramifications of differences between twin funds, these three reasons identify possible channels through which institutional investor monitoring of one portfolio could affect a twin retail portfolio, namely, direct expenses, indirect expenses and manager effort. In Table 8 we explore these channels by analyzing changes in the net expense ratios, return gap, active share and other relevant fund characteristics, before and after the addition of an institutional twin. The first row explores change in a fund's direct expenses, namely the expense ratio, before and after the creation of an institutional twin. The data for this analysis is taken from a database of SEC N-SAR filings that is described in Edelen, Evans and Kadlec (2011a). The expense ratio of the treatment group decreases after the creation of the institutional twin, by 5.3 basis points on average. Because the expense ratio of the control group increases over the same time period by 0.3 basis points, we find a total control-group adjusted change in the net expense ratio of 5.6 basis points.[24]

Separate from the direct expenses, indirect costs that are not included in the expense ratio such as brokerage commissions, trading/implementation costs, and the like, could also affect performance. There is evidence that indirect expenses are opaque to retail investors (e.g. Edelen, Evans and Kadlec (2011a)), but if institutional investors are aware of them, monitoring of those expenses to ensure they are not excessive could benefit retail investors. To measure these indirect costs, we use the return gap measure of Kacperczyk, Sialm and Zheng (2008). The return gap is the difference between the actual fund return and the return of the fund based on the previously disclosed holdings minus the fund's expense ratio. By comparing the actual fund return to the return on a hypothetical portfolio return constructed from a buy

[24] We report results on the expense ratio using the NSAR data for consistency because brokerage commissions and soft dollar distributions are calculated from NSAR as well. Using Morningstar data to calculate the levels and changes in expense ratios increases the sample to 123 funds, and leads to a diff-in-diff of 3.3 basis points, statistically significantly different from zero at the 1% level.

and hold strategy of the fund's last disclosed holdings, the return gap measures the aggregate value added (positive return gap) or destroyed (negative return gap) by a manager's actions above and beyond the fund's expense ratio. The second row of Table 8 examines the return gap of the treatment and control fund samples. It shows that the annualized return gap is positive for the treatment sample, i.e. these funds on average provide greater hidden benefits than costs, but negative for the control sample. There is also some evidence that fund advisors choose funds with a favorable return gap when they add an institutional twin: Prior to the creation of the institutional twin, the difference in return gap between treatment and control group is a positive 0.321% annually. More importantly, we see substantial improvement in the return gap (i.e., greater outperformance relative to the return on previously disclosed holdings) which almost doubles after the institutional twin creation for the treatment sample. Finally, we observe deterioration in the return gap for the control sample. The diff-in-diff effect is an economically large improvement in the return gap measure of 0.552% annualized. It is important to note that in addition to measuring indirect costs, the return gap is also a potential measure of manager effort/skill, capturing short term performance enhancement/degradation due to trading. Consequently, the return gap result is also consistent with increased manager effort translating to improved fund performance.

The return gap is an aggregate of many unobserved managerial actions that are difficult to directly measure. However, we have data for two components of the return gap, brokerage commissions and soft dollar usage, and in Table 8 we examine how these components change before and after the addition of an institutional twin. Both brokerage commissions (Q21) and the use of soft dollars to pay for distribution (Q26.A) variables come from the N-SAR filings. Edelen, Evans and Kadlec (2011a) show that both of these costs are strongly negatively related to future performance and that they are opaque to retail investors. Recognizing their impact on performance, an institutional investor with greater awareness of these costs might discourage their use by the investment advisor, and the investment advisor, for reasons outlined at the beginning of the section, might implement these changes for retail funds too. Looking at the results in row 4 of Table 8, we do not find a statistically significant change in brokerage commission rates. However, after the creation of a twin, we observe a substantial decrease of

5.8% of retail mutual funds that use soft dollar distribution. Because the fraction of funds using soft dollar distributions decreases for the control group as well, we find and overall highly statistically significant diff-in-diff effect of -3.6%. Relative to the fraction of funds using soft-dollar distribution prior to the creation of the twin fund, the decrease is 15.5% and appears economically significant.

In addition to decreasing fees and increasing the return gap, performance could also be improved through increased managerial effort resulting in superior stock selection. Because the return gap compares the actual fund return to the buy and hold return on the previously disclosed portfolio holdings (i.e. the benchmark for the manager is the manager's own previously disclosed portfolio), it abstracts from the value added through a manager's superior stock selection.[25] While it is difficult to directly measure managerial effort, we attempt to proxy for it by examining portfolio characteristics before and after the institutional twin is added as a proxy. First, we examine Active Share, a measure of the overlap between the fund's holdings and the closest related index developed by Petajisto (2010) and Cremers and Petajisto (2009). Cremers and Petajisto (2009) construct Active Share as the percentage of portfolio holdings that differ from the benchmark index holdings. For mutual funds, active share typically varies between 0% (pure index funds) and 100%, and Cremers and Petajisto (2009) report an average active share of approximately 63% in 2002. They show that funds with the highest Active Share significantly outperform their benchmarks, both before and after expenses. Our second proxy is the value-weighted average number of analyst estimates. It attempts to capture whether fund managers invest in stocks with greater analyst coverage, an indication of less effort, or less analyst coverage, consistent with greater effort.

Row 5 of Table 8 shows the results for Active Share, measured over the 36 months before and after the creation of the institutional twin.[26] Prior to the creation of the institutional twin, treatment and

[25] It does include the value added through the manager's intra-holding period trades. In contrast, the performance measurement methodology used in Daniel, Grinblatt, Titman and Wermers (1997) and Wermers (2000) uses the return on a portfolio of characteristic-based stocks as the benchmark and not the manager's portfolio itself.

[26] We measure Active Share over a three-year horizon, because the asset allocation of funds may be difficult to change short-term. When we measure changes in Active Share annually to examine robustness of this result, we find

control group both have an Active Share of between 76 and 77%, which is statistically indistinguishable from each other. The change in Active Share for the treatment group is statistically indistinguishable from zero, while the change in Active Share for the control group has a negative and statistically significant point estimate of approximately 1% (i.e. more overlap with the benchmark or less active stock selection on the part of the manager). The difference-in-difference result shows that there is a statistically significant increase in Active Share for the treatment group relative to the control group.[27]

Row 6 of Table 8 shows that, prior to the event, treatment firms hold, on average, stocks with more analyst coverage. After the institutional twin is added, on average the treatment group holds stocks with slightly lower analyst coverage and the control group holds stocks with greater analyst coverage, showing a significant decrease in the value-weighted number of analysts of the treatment sample relative to the control. While the changes in analyst coverage are relatively small, this evidence is consistent with managers making a greater effort to identify stocks.

We repeat the propensity score analysis of the variables in Table 8 with different caliper lengths (0.001 and 0.005) and different number of maximum matches (5 and 10), as in Table 7, to ensure the robustness of our results. Our results are quantitatively and qualitatively similar to those reported in Table 8.

To summarize, while our matched sample analysis shows that adding an institutional twin is associated with decreased direct expenses for retail mutual funds, the effect is small (0.056% annually) relative to the overall impact on annual performance we measured using the same sample in Section 3.3 (1.548%). The fee analysis compares the first year before and after the addition of a twin and the performance results compare the three years before and after, and fees might continue to decrease in years two and three. However, it seems unlikely they would decrease enough in those additional years to account for the full 1.548% performance difference. In contrast, the return gap analysis shows a

---

a diff-in-diff result that is approximately half of the reported result, and statistically significantly different from zero at the 10% level.

[27] Given the passive nature of index funds, including them in our analysis may weaken the difference results. In unreported tests, we repeat the active share analysis after removing index funds from the treatment and control samples and we find that the improvement in active share is a statistically significant 0.89%.

performance improvement of 0.552% after adding an institutional twin, just over a third of the performance improvement. Separate from direct expenses and the indirect expenses measured by the return gap, increased managerial effort in selecting equities in the portfolio as a result of the increased monitoring could also improve performance. While we cannot measure increased effort and any resulting performance improvement directly, our results using proxies for managerial effort are consistent with this interpretation. Overall, our analysis suggests some channels through which institutional investor monitoring could improve performance in twin retail mutual funds.

5. Robustness

We demonstrate in Section 3 that retail mutual fund performance improves after the creation of an institutional fund twin. Relative to a carefully chosen control group, retail fund performance increases and fees decrease after the creation of the twin. We argue in Section 4 that this effect is consistent with increased monitoring of the fund by institutional investors. To provide further support for our hypothesis, in section 5.1 we examine the results of a placebo experiment and in section 5.2 we repeat the flow and performance analyses separating out institutional share classes.

5.1. Placebo experiment

Our sample of 463 twin funds also contains 179 twins in which the retail fund was created *after* the institutional fund. Given our hypothesis of provision of greater market governance by institutional investors relative to retail investors, we would not expect to observe market governance driven increases in performance in institutional funds after the creation of retail funds. If we were to observe performance regression results for institutional funds similar to those reported it Table 5 or diff-in-diff results similar to those reported in Table 7, we would have to reject the monitoring hypothesis in favor of a hypothesis that other factors in the contractual environment (e.g., larger fund, economies of scale related to fees) are responsible for the reported changes.

We analyze our placebo experiment using both the Heckman two-stage approach used in Table 5 and the propensity score matched sample in Table 6 and 7. For the Heckman regression, Panel A of Table 9 gives the results from the first stage probit regression in which the dependent variable is equal to one if the institutional fund has a retail twin in month $t-1$, and zero otherwise. Panel B contains the results from the second-stage performance regression where a Heckman correction term is included.

Panel A shows that the probability that a retail twin fund exists in any given month is increasing in the past performance of the institutional fund as measured by the 4-factor alpha over the past 36 months and the size of the institutional fund. The indicator for whether or not the fund family has any other retail funds is also positive, indicating that families that already have retail fund offerings are more likely to create a retail twin for one of their institutional offerings.

Panel B shows the results of the monthly performance regressions that now includes the Heckman correction term (lambda). Consistent with our evidence on the monitoring value of institutional investors relative to retail investors, there is no statistically significant performance improvement after a retail twin is added. Comparing the other coefficients with the retail fund regression results in Panel B of Table 5, the biggest difference is the insignificant coefficient on institutional family size in specification 1. As we explained earlier, because the Morningstar database does not capture all institutional assets, our measure of family institutional TNA may be understating true family size and as a result it may not be capturing potential economies of scale for families with a larger institutional presence.

In addition to the Heckman regressions, we also repeat our placebo analysis using a propensity score matching approach. To set up the propensity matched placebo test, we first estimate a probit model predicting which institutional funds will create a twin retail fund, using characteristics similar to those used in Panel A of Table 6.[28] The estimates from this probit model are given in Panel A of Table 10. In Panel B we also give a comparison of the treatment and control sample characteristics to assess the

[28] While the probit model in Table 4 also included fund turnover and an indicator variable of whether or not the fund was sold by a broker, these variables are excluded in the model in Table 7. The institutional data does not provide information about broker distribution and including the turnover variable does not affect the results but it reduces the sample size significantly.

quality of the matching between them. The main drivers of the decision to create a retail twin fund are high past performance, large retail family TNA, and a high expense ratio. Panel B of Table 10 shows that the matching works well. The means and medians of fund characteristics for treatment and control group are statistically indistinguishable from each other with the exception of 4-factor alpha. However, given the lack of performance persistence documented by Busse, Goyal and Wahal (2010) for institutional investment funds, this difference in the treatment and control sample should not bias our results.[29]

Using this matched sample, we examine the addition of a retail twin fund and the corresponding change in the risk-adjusted performance of the institutional fund relative to a control sample, similar to Table 7. The results of this analysis are included in Table 11. We see that the treatment sample, i.e., those institutional twin funds that add a retail twin, had risk-adjusted performance of 3.28% per year prior to the event. Given the focus of retail investors on past good performance that we examined in Table 3 and the ability of fund sponsors to use the performance of the institutional predecessor account in advertising the retail twin with certain required disclosures[30] it is perhaps not surprising that the treatment sample outperforms the control sample by a statistically significant 1.12% per year. Consistent with diseconomies of scale documented by Chen, Hong, Huang and Kubik (2004), both treatment and control funds exhibit significant declines in their risk-adjusted performance in the 36 months after. In the three years after adding a retail twin, the difference in performance between the treatment and control funds is a statistically insignificant 0.112% per year. Overall, the diff-in-diff results show that the change in performance of institutional funds with added retail twins compared to the change in performance of the control group is negative. This result stands in sharp contrast to the performance improvement of retail funds observed in Table 7 and further support the role of institutional investor monitoring as an

[29] We also repeat the analysis including only the 4-factor alpha in the propensity score probit model. By only including this variable in the propensity score model we eliminate the difference in the treatment and control sample 4-factor alphas, but the diff-in-diff performance results are largely unchanged. These results are available upon request.

[30] Pierce (1998) outlines the SEC's criteria for allowing mutual fund sponsors to adopt and advertise the performance record of an unregistered predecessor account.

explanation for our results.[31] As we discussed in Section 3.1, separate accounts are not subject to the same reporting requirements as mutual funds. As a result, the separate account data in the Morningstar database is provided on a voluntary basis and may be subject to both survivorship and backfill biases. Using the filters described in Section 3.1, we repeat our analysis for a backfill-free and survivorship bias free sample and the results are given in row 2 of Table 11. With the filtered data, the performance in the 36 months before the retail twin creation is substantially lower and there is no difference between the treatment and control. However, consistent with the observed increase in assets under management, both the treatment and control groups experience decreased performance in the 36 months after and we find no statistically significantly difference in performance between the treatment and control group.

5.2. Institutional share classes

In addition to the placebo test described in section 5.1, institutional share classes provide us with another test of the potential role of institutional investors as monitors. To this point in the paper, we have categorized mutual funds that have only retail share classes and mutual funds that have both institutional share classes and retail share classes into one group that we term "retail" in our main empirical analysis. While our more cautious classification should bias us against finding any difference between retail funds with and without twins, one could argue that a mutual fund that has significant assets in institutional share classes would also benefit from institutional oversight and thus should be classified as a retail fund with an institutional twin. At the same time, anecdotal evidence[32] suggests that investors in institutional share

[31] Because the separate account institutional twins are not subject to the Investment Company Act of 1940 and are therefore not required to report N-SARs, we don't have access to the variables necessary to repeat the analysis of Table 8 for our placebo sample.

[32] For example, a recent Harvard Business School case study on the Yale Investment Office describes: "Yale's commitment to this asset class was tested in 1998, when many hedge funds suffered in the « flight to liquidity » that followed Russia's August 1998 default on its debt obligations. […] Even though some of these pricing anomalies were likely to be short-lived […] a number of investors panicked and demanded the return of their capital. As a result, some funds were forced to liquidate positions at exceedingly unfavorable prices. While in most cases the university was insulated from the effects of other investors' sales because the fund managers had established separate accounts for Yale's investment, in other cases, Yale's funds were commingled with those of other investors. In these instances, Yale's returns suffered. As a result of this experience, Yale redoubled its efforts to utilize separate accounts…" [Yale University Investments Office: August 2006, Harvard Business School Case No. 807-073].

classes may not be as sophisticated as separate account or pure institutional mutual fund investors. A separate institutional fund or a separate account has the advantage that it shields the institution's investment from potential trading inefficiencies created by retail investor flow volatility. In addition, management fees for these institutional vehicles are typically lower than for their retail counterparts, even the institutional share classes. Large, sophisticated institutional investors, which we think are the most likely to provide oversight, would be aware of these advantages and we would expect they would opt for a separate account or a separate investment vehicle instead of the institutional share class. Hence, we would expect our results on market governance to be stronger for separate account and institutional fund twins than for institutional share class twins.

We repeat the analysis of flow determinants in Table 3, separating out flows to retail funds/share classes, institutional share classes, and separate accounts/institutional funds, to provide further evidence for this conjecture. To do this we identify any fund that has both retail and institutional share classes. We then aggregate the variables for the retail share classes for that fund each month into a single retail fund-month observation and similarly for the institutional share classes and treat these two fund-month observations as a retail-institutional twin fund pair. The results of this flow analysis are included in Table 12. Comparing the institutional share class coefficient estimates to the retail we find that while institutional share class flows are more sensitive to expense ratios than the retail flows, they exhibit little sensitivity to poor risk-adjusted performance and are not statistically or economically different from retail flows in this regard. Additionally, the institutional share class flows are actually more sensitive to total returns both for the high and low performance than the retail flows. This is consistent with institutional investors in these institutional share classes being less sophisticated than institutional investors in separate accounts or institutional funds.

Given the institutional share class flow results, we would expect that institutional share class twins do not provide as effective monitoring as the separate account and institutional mutual fund twins. As a result, we would expect the existence of an institutional share class twin to have less of an effect on performance than the other institutional twin funds. To test this, we repeat the performance regressions in

Table 4, separating out twin identifiers for funds with separate account/institutional fund twins and institutional share class twins. The results of this analysis are included in Table 13. Consistent with less sophisticated investor monitoring in the institutional share classes, we find no statistically significant improvement in performance for retail funds after the creation of an institutional share class. The outperformance of retail twin funds before the creation of an institutional share class is not surprising given the flow evidence that institutional share class investors are responsive to the good risk-adjusted performance. As a result, an advisor would be more likely to offer an institutional share class for retail funds with high alpha.

6. Conclusion

The Investment Company Act of 1940 gave investors in open-end mutual funds a unique and innovative governance mechanism – the ability to redeem. Because the decision to redeem shares and the associated loss of management's control over these assets can be undertaken independently by each investor no matter how small, they can effectively remove the fund manager from the control of those assets. Recognizing the significant role played by "market governance", Fama and Jensen (1983) suggest that it is primary to all other fund governance mechanisms. The effectiveness of this governance mechanism in protecting shareholders, however, depends on how investors exercise their right to redeem and whether or not they respond to useful investment signals such as fees and poor past risk-adjusted performance.

Using a sample of retail mutual funds with an institutional twin, a similar but separately managed institutional fund, we examine how retail and institutional investors in similar investment products respond to these investment signals. We find that institutional investor flows are more sensitive to high fees and to poor risk-adjusted performance than retail flows. Additionally, retail investors respond more strongly to counterproductive signals like past total return than institutional investors.

We also examine the relationship between the creation of an institutional twin the performance of its retail fund counterpart. Consistent with greater monitoring on the part of the institutional twin investors, retail fund risk-adjusted performance increases by 1.5% per year if the fund manager also manages an institutional twin. Exploiting cross-sectional differences in the date the institutional and retail mutual funds were created, we are able to examine whether institutional investors are merely better at selecting managers, or whether their presence reduces agency problems between mutual fund managers and investors. We uncover evidence consistent with the latter explanation. Fees, trading costs, and other fund expenses decrease and managerial effort increases relative to a control sample after the institutional twin is created.

Our results have important implications for recent legal and legislative developments related to excessive mutual fund fees. In 2003, Eliot Spitzer, then the New York Attorney General, criticized investment advisers in the mutual fund industry for setting higher advisory fees for retail mutual fund investors than for clients in corresponding institutional separate accounts. More recently, a comparison of the fees between retail and institutional twin funds was at the heart of the Jones v. Harris Associates appeal heard by the Supreme Court in 2010. Indeed, the premise behind much of the excessive fee litigation has been a breach of fiduciary duty on the part of advisors because they charge 'excessively' high fees to retail clients relative to what they charge their separate account clients for the same or a very similar investment product. While these cases have so far been rejected by courts, they have created some legal uncertainty for fund families that pursue twin arrangements. Some commentators have gone so far as to speculate whether mutual fund families will shut down and avoid such twin fund arrangements in the future to avoid litigation.

In contrast with the negative view suggested by these events, our evidence provides a positive perspective on differences in retail and institutional fund fees. If twin fund arrangements indeed provide better monitoring it would be a justification for these fee differentials. In this light, the differential may be viewed as analogous to a firm paying a director in return for the monitoring provided. The average expense ratio difference between institutional and retail twins in our sample is 0.42% and significantly smaller than the risk-adjusted excess performance of 1.5% of those retail funds with institutional twins. Our evidence regarding the improved performance and manager effort and the reduced direct and indirect expenses of retail twin funds suggests that monitoring by investors in institutional twins serves as an important governance mechanism for investors in retail twins.

References

Adams, John C., Sattar A. Mansi, and Takeshi Nishikawa, 2010, Internal governance mechanisms and operational performance: Evidence from index mutual funds, *Review of Financial Studies* 23, 1261-1286.

Aggarwal, Reena, Isil Erel, René M. Stulz, and Rohan Williamson, 2009, Differences in governance practices between the U.S. and foreign firms: Measurement, causes, and consequences, *Review of Financial Studies* 22, 3131-3169.

Agrawal, Anup, and Gershon N. Mandelker, 1990, Large shareholders and the monitoring of managers: The Case of anti-takeover charter amendments, *Journal of Financial and Quantitative Analysis* 25, 143-161.

Almazan, Andres, Jay C. Hartzell, and Laura T. Starks 2005, Active institutional shareholders and managerial compensation, *Financial Management* 34, 5-34.

Berk, Jonathan B. and Richard C. Green, 2004, Mutual fund flows and performance in rational markets, *Journal of Political Economy* 112, 1269-1295.

Brown, Keith, W.V. Harlow, and Laura Starks, 1996, Of tournaments and temptations: An analysis of managerial incentives in the mutual fund industry. *Journal of Finance* 51, 85-110.

Busse, Jeffrey A., Amit Goyal and Sunil Wahal, 2010, Performance and persistence in institutional investment management, *Journal of Finance* 65, 765-790.

Carhart, Mark, 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 1, 57-82.

Chen, Joseph, Harrison G. Hong, Ming Huang, and Jeffrey D. Kubik, 2004, Does fund size erode mutual fund performance? The role of liquidity and organization, *American Economic Review* 94, 1276-1302.

Chen, Qi, Itay Goldstein, and Wei Jiang, 2008, Directors' ownership in the U.S. mutual fund industry, *Journal of Finance* 63, 2629-2677.

Chevalier, Judith and Glenn Ellison, 1997. Risk taking by mutual funds as a response to incentives. *Journal of Political Economy* 105, 1167-1200.

Cici Gjergji, Scott Gibson, and Rabih Moussawi, 2010, Mutual fund performance when parent firms simultaneously manage hedge funds, *Journal of Financial Intermediation* 19, 169-187.

Cremers, Martijn, Joost Driessen, Pascal J. Maenhout, and David Weinbaum, 2009, Does skin in the game matter? Director incentives and governance in the mutual fund industry, *Journal of Financial and Quantitative Analysis* 44, 1345-1373.

Cremers, Martijn, and Antti Petajisto, 2009, How active is your fund manager? A new measure that predicts performance, *Review of Financial Studies* 22, 3329-3365.

Daniel, Kent, Mark Grinblatt, Sheridan Titman and Russ Wermers, 1997, Measuring mutual fund performance with characteristic-based benchmarks, *Journal of Finance* 52, 1035-1058.

Dann, Larry, Diane Del Guercio and Megan Partch, 2003, Governance and boards of directors in closed-end investment companies, *Journal of Financial Economics* 69, 111-152.

Del Guercio, Diane, Laura Seery, and Tracie Woidtke, 2008, Do boards pay attention when institutional investor activists just vote no, *Journal of Financial Economics* 90, 84-103.

Del Guercio, Diane and Paula A. Tkac, 2002, The determinants of the flow of funds of managed portfolios: Mutual funds vs. pension funds, *Journal of Financial and Quantitative Analysis* 37, 523-557.

Ding, Bill and Russ R. Wermers, 2005, Mutual fund performance and governance structure: The role of portfolio managers and boards of directors, *Working Paper, University of Maryland.*

Drucker, Steven and Manju Puri, 2005, On the benefits of concurrent lending and underwriting, *Journal of Finance* 60, 2763–2799.

Edelen, Roger M., Richard B. Evans, Gregory B. Kadlec, 2011a, Disclosure and agency conflict in delegated investment management: Evidence from mutual fund commission bundling, *Journal of Financial Economics*, forthcoming.

Edelen, Roger M., Richard B. Evans, Gregory B. Kadlec, 2011b, Information and implementation: Assessing the net impact of trading on mutual funds, *Working Paper, University of Virginia.*

Fama, Eugene F. and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F. and Michael C. Jensen, 1983, Separation of ownership and control, *Journal of Law and Economics* 26, 301-325.

Gillan, Stuart, and Laura Starks, 2000, Corporate governance proposals and shareholder activism: The role of institutional investors, *Journal of Financial Economics* 57, 275-305.

Goyal, Amit and Sunil Wahal, 2008, The selection and termination of investment management firms by plan sponsors, *Journal of Finance* 63, 1805-1847.

Greene, William H., 2011, Econometric Analysis, Prentice Hall (7th Edition), Upper Saddle River, NJ.

Hartzell, Jay C., and Laura T. Starks, 2003, Institutional investors and executive compensation, *Journal of Finance* 58, 2351-2374.

Heckman, James, Hidehiko Ichimura, and Petra Todd, 1997, Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme, *Review of Economic Studies* 64, 605-654.

Heckman, James, Hidehiko Ichimura, and Petra Todd, 1998, Matching as an econometric evaluation estimator, *Review of Economic Studies* 65, 261-294.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2008, Unobserved actions of mutual funds, *Review of Financial Studies* 21, 2379-2416.

Khorana, Ajay, and Henri Servaes, 1999, The Determinants of Mutual Fund Starts, *Review of Financial Studies* 12, 1043-1074.

Khorana, Ajay, Peter Tufano, and Lei Wedge, 2007, Board structure, mergers and shareholder wealth: A study of the mutual fund industry, *Journal of Financial Economics* 85, 571-598.

Maddala, Gangadharrao S., 1983, Limited-Dependent and Qualitative Variables in Econometrics, Cambridge University Press, Cambridge, MA.

Meschke, J. Felix, 2008, An empirical examination of mutual fund boards, *Working Paper, University of Minnesota.*

Morley, John and Quinn Curtis, 2010, Taking exit rights seriously: Why governance and fee litigation don't work in mutual funds, *Yale Law Journal* 120, 84-142.

Newey, Whitney K. and Kenneth D. West, 1987, A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* 55, 703–708.

Nohel, Tom, Z. Jay Wang, and Lu Zheng, 2010, Side-by-side management of hedge funds and mutual funds, *Review of Financial Studies* 23, 2342-2373.

Parrino, Robert, Richard W. Sias, and Laura T. Starks, 2003, Voting with their feet: Institutional ownership changes around forced CEO turnover, *Journal of Financial Economics* 68, 3-46.

Petajisto, Antti, 2010, Active share and mutual fund performance, *Working Paper, Yale School of Management.*

Pierce, Leonard A., 1998, Portability of performance records and the use of related performance information, *Journal of Performance Measurement* 3, 22-34.

Rosenbaum, Paul R., and Donald B. Rubin, 1983, The central role of the propensity score in observational studies for causal effects, *Biometrika* 70, 41–55.

Sirri, Erik R. and Peter Tufano, 1998, Costly search and mutual fund flows, *Journal of Finance* 53, 1589-1622.

Treynor, Jack L. and Fisher Black, 1973, How to use security analysis to improve portfolio selection, *Journal of Business* 46, 66–88.

Tufano, Peter, and Matthew Sevick, 2010, Board structure and fee-setting in the U.S. mutual fund industry, *Journal of Financial Economics* 46, 321-356.

Wermers, Russ, 2000, Mutual fund performance: An empirical decomposition into stock-picking talent, style, transactions costs, and expenses, *Journal of Finance* 55, 1655-1695.

White, Halbert, 1980, A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica* 48, 817–838.

Yan, Xuemin, 2008, Liquidity, investment style, and the relation between fund size and performance, *Journal of Financial and Quantitative Analysis* 43, 741-768.

Zitzewitz, Eric, 2003, Who cares about shareholders: Arbitrage-proofing mutual funds, *Journal of Law, Economics, and Organization* 19, 245-280.

Table 1 – Retail mutual fund sample summary statistics

Table 1 provides descriptive statistics for the sample of US domestic equity mutual funds from 1997 to 2009 used in the paper. Panel A lists the number of fund-month observations in the sample by year. Panel B lists the number of funds, and percent of the sample, with and without an institutional twin. The funds with twins are further divided into those with a separate account twin versus those with an institutional mutual fund twin. Panel C gives the mean, median, 25th and 75th percentiles for fund and fund family characteristics, separating the sample of fund-month observations into those with and without institutional twins. The descriptive statistics are given for fund and family total net assets (TNA) under management in millions ($MM) and billions ($BB) of dollars respectively. It also lists the annual expense ratio charged to investors in the fund, the fund's turnover (minimum of the fund's purchases and sales for the 12 month period divided by average fund TNA over the same period), quarterly net flow as a percent of TNA, and the fund's annualized four-factor alpha (the three Fama and French (1993) factors combined with the momentum factor of Carhart (1997)), calculated over the previous 36 months. It also lists the percentage of broker-sold fund-month observations, where broker-sold is defined as any fund that has a front load, a rear load or charges a 12b-1 fee greater than 0.25%. For the twins sample, Panel C also provides the relative start date of the twin. This is calculated as the inception date of the institutional twin minus the inception date of the retail twin in years. Panel D breaks the sample of retail-institutional matched twins in three groups: those where the retail fund was started first followed by the institutional twin, those where the institutional twin was started first and those where they started at the same time. Panel D lists the number of twins in each category for the full sample and for the subsample where the inception date of the relevant fund occurs after the start of our sample period, 1996 (institutional twin inception date if it is in the retail fund first category; retail twin inception date if it is in the institutional fund first category; and either if it is in the same inception date category).

Panel A. Observations by year

Year	1997	1998	1999	2000	2001	2002	2003
Observations	8,543	9,597	10,436	12,704	14,405	15,910	16,994
Funds	760	847	963	1,159	1,285	1,402	1,569

Year	2004	2005	2006	2007	2008	2009
Observations	18,497	20,567	20,598	21,603	22,155	22,351
Funds	1,640	1,842	1,775	1,911	1,964	1,944

Panel B. Sample composition

	No. of Funds	Percent
No Twin	2,197	82.6%
Twin	463	17.4%
Separate Account Twin	345	13.0%
Institutional MF Twin	118	4.4%

Table 1 – Retail mutual fund sample summary statistics (Continued)

Panel C. Univariate statistics

No Twins Sample (172,846 Obs.) Variable	Mean	Median	Percentiles 25th	75th
Family TNA ($BB)	31.9	5.3	0.7	23.9
Fund TNA ($MM)	1,496.5	241.4	66.4	874.3
Expense Ratio (Annual)	1.27%	1.25%	0.96%	1.54%
Turnover (%)	91%	61%	29%	110%
Quarterly Net Flow (%)	2.35%	-0.50%	-3.82%	4.24%
4-Factor Alpha (Annualized)	-1.46%	-1.56%	-12.35%	9.11%
Broker-Sold	64.1%			

Twins Sample (41,514 Obs.) Variable	Mean	Median	Percentiles 25th	75th
Family TNA ($BB)	24.6	8.8	2.3	33.7
Fund TNA ($MM)	1,423.2	412.4	126.8	1,297.9
Expense Ratio (Annual)	1.24%	1.23%	0.97%	1.53%
Turnover (%)	86%	65%	34%	111%
Quarterly Net Flow (%)	3.31%	-0.02%	-3.45%	5.41%
4-Factor Alpha (Annualized)	-0.75%	-1.00%	-11.69%	9.95%
Relative Start Date (Inst. Incept. Date - Retail Incept. Date)	0.9	0.0	-3.5	6.1
Broker-Sold	67.3%			

Panel D. Twin start dates

	Full Sample			Inception Date between 1996 and 2009		
	Retail Fund First	Institutional Fund First	Same Incept. Date	Retail Fund First	Institutional Fund First	Same Incept. Date
All Twins	237	224	2	132	179	2

Table 2 – Institutional fund sample summary statistics

Table 2 provides descriptive statistics for our sample of institutional fund twins from 1997 to 2009. Panel A gives the mean, median, 25th and 75th percentiles and number of observations for fund and family total net assets (TNA) under management in millions ($MM) and billions ($BB) of dollars respectively. It also lists the annual expense ratio charged to investors in the fund, the fund's turnover (minimum of the fund's purchases and sales for the 12 month period divided by average fund TNA over the same period), quarterly net flow as a percent of TNA, and the fund's annualized four-factor alpha (the three Fama and French (1993) factors combined with the momentum factor of Carhart (1997)), calculated over the previous 36 months. Panel B contains summary statistics for the differences in characteristics between the retail mutual funds and their institutional twins for the above variables and, in addition, for differences in annualized gross and net total return and differences in the factor loadings on the three Fama and French (1993) factors (market (MKT), value (HML), and size (SMB) and the Carhart (1997) momentum factor. Panel B provides the mean, median, 25th and 75th percentiles as well as p-values from a t-test of the difference in means and a sign test of the difference in medians. The monthly retail fund data is merged with the corresponding monthly institutional twin data. For each matched pair-month observation, the difference in the variables of interest is calculated. These differences are averaged for each of the matched pairs and cross-sectional sample statistics from the matched pairs are given. The last column shows the number of twin pairs with complete data for each variable.

Panel A. Institutional twin sample

Variable	Mean	Median	Percentiles 25th	75th	Obs.
Family TNA ($BB)	33.5	10.8	3.1	53.9	6,200
Fund TNA ($MM)	1,788.8	694.8	174.5	1,811.0	6,200
Expense Ratio	0.80%	0.72%	0.50%	0.96%	17,919
Turnover (%)	80%	63%	34%	106%	10,694
Quarterly Net Flow (%)	3.51%	-0.32%	-3.65%	4.61%	4,335
4-Factor Alpha (Annualized)	-0.56%	-0.72%	-11.46%	10.10%	18,205

Panel B. Average fund differences

Variable	Diff. (Inst - Retail) Mean	Median	Percentiles 25th	75th	Diff. Test p-Values Mean	Median	Number of Twin Pairs
Fund TNA ($MM)	547.9	86.8	-209.8	816.4	0.050	<.001	435
Expense Ratio (Annual)	-0.42%	-0.44%	-0.73%	-0.16%	<.001	<.001	437
Turnover (%)	-7.3%	-0.4%	-13.7%	6.7%	0.004	0.147	298
Quarterly Net Flow (%)	2.04%	0.12%	-4.76%	6.94%	0.079	0.460	310
4-Factor Alpha (Annualized)	0.54%	0.39%	-0.10%	0.99%	<.001	<.001	420
Net Total Return (Annualized)	0.62%	0.51%	-0.02%	1.17%	<.001	<.001	463
Gross Total Return (Annualized)	0.13%	-0.01%	-0.43%	0.45%	0.047	0.376	463
Pct in Common Stock (%)	3.5%	3.3%	0.0%	7.7%	<.001	<.001	367
MKT Factor (x100)	-0.087	0.039	-1.114	1.051	0.666	0.354	420
HML Factor (x100)	0.389	0.025	-0.935	1.177	0.138	0.884	420
SMB Factor (x100)	-0.476	-0.036	-1.575	0.952	0.110	0.354	420
Momentum Factor (x100)	-0.107	-0.042	-0.761	0.546	0.472	0.157	420

Table 3 – Determinants of institutional vs. retail flows

Table 3 presents coefficient estimates and t-statistics from pooled regressions of quarterly net fund flow on lagged fund characteristics. The first specification is estimated using a sample that includes all retail and institutional domestic equity funds over the period 1997 through 2009. The second specification is estimated using a sample that only includes retail and institutional funds from the matched twin sample over the period 1997 through 2009. The dependent variable is the quarterly fund flow (t=0 to t=3) as a percent of TNA. The independent variables include an intercept, the natural log of lagged (t=-1) fund family and fund TNA, the lagged fund expense ratio, lagged turnover, the concurrent (t=0 to t=3) percentage quarterly flow to funds with the same investment objective (net of own fund flow), the lagged percentage quarterly fund flow, the concurrent quarterly flow in the twin fund (equal to zero for all non-twin funds), and two different measures of performance: 36-month total return and 4-factor alpha (Fama and French (1993); Carhart (1997)), calculated over the previous 36 months. For both performance measures, we allow for non-linearity through the use of a piece-wise linear performance specification with kinks at the 20th and 80th percentile of returns. For the total return measure, the percentiles are calculated within date and investment objective similar to Sirri and Tufano (1998), while the 4-factor alpha percentiles are calculated within date only. With these percentiles, the formula for the low return is LowRet=Min(0.2, RetPtile), the formula for the medium return is MedRet=Min(0.6, RetPtile-LowRet) , and the formula for the high return is HighRet=RetPtile-MedRet-LowRet. Specifications 1 and 2 allow for separate coefficients for the retail and institutional funds. For both specifications, p-values from difference in coefficients tests across the retail and institutional coefficients for the expense ratio, total return, and 4-factor alpha are provided at the bottom of the table. Standard errors are clustered by fund and by date and the total number of quarterly fund observations and the adjusted R-squared are provided.

Regression:	1				2			
	Retail Coef.		***Inst. Coef.***		***Retail Coef.***		***Inst. Coef.***	
Intercept	0.1203	(6.8)	0.3701	(4.5)	0.1418	(5.1)	0.4324	(2.8)
Log(Family TNA)$_{t-1}$	0.0051	(9.0)	0.0047	(2.9)	0.0044	(3.6)	0.0005	(0.1)
Log(Fund TNA)$_{t-1}$	-0.0144	(-15.0)	-0.0283	(-9.1)	-0.0146	(-8.1)	-0.0254	(-5.9)
Expense Ratio$_{t-1}$	-0.0079	(-4.3)	-0.0468	(-3.5)	-0.0063	(-1.5)	-0.0600	(-3.0)
Turnover$_{t-1}$	0.0026	(1.6)	-0.0031	(-0.7)	0.0008	(0.6)	-0.0008	(-0.3)
InvObj Qtrly Pct Flow$_{t,t+3}$	0.4145	(8.4)	0.9603	(7.0)	0.3744	(5.1)	0.8998	(3.8)
Qtrly Pct Flow$_{t-4,t-1}$	0.2432	(15.1)	0.0130	(1.0)	0.3125	(10.5)	0.0148	(0.4)
Twin Fund Qtrly Pct Flow$_{t,t+3}$	0.0294	(3.3)			0.0277	(3.2)		
Total Return$_{t-36,t-1}$ Low (≤ 20th Ptile)	0.0587	(3.0)	0.1621	(3.2)	0.1368	(3.1)	0.0347	(0.3)
Total Return$_{t-36,t-1}$ Medium (> 20th Ptile & ≤ 80th Pt	0.0431	(9.2)	0.0104	(0.6)	0.0388	(4.8)	-0.0384	(-1.4)
Total Return$_{t-36,t-1}$ High (> 80th Ptile)	0.1987	(6.0)	0.0444	(0.7)	0.1097	(2.0)	-0.0634	(-0.4)
4-Fctr Alpha$_{t-36,t-1}$ Low (≤ 20th Ptile)	0.0890	(5.3)	0.2171	(4.4)	0.0599	(1.4)	0.3095	(2.8)
4-Fctr Alpha$_{t-36,t-1}$ Medium (> 20th Ptile & ≤ 80th Pt	0.0281	(6.4)	0.0826	(4.9)	0.0253	(2.9)	0.1043	(3.5)
4-Fctr Alpha$_{t-36,t-1}$ High (> 80th Ptile)	0.1876	(5.2)	0.1945	(3.4)	0.1416	(2.5)	0.2651	(2.0)
Observations	108,803				21,733			
Adj. R-Squared	8.25%				10.07%			
Coef. Difference Test p-Values (Retail vs. Institutional)								
Expense Ratio$_{t-1}$	0.005				0.009			
4-Fctr Alpha$_{t-36,t-1}$ Low (≤ 20th Ptile)	0.014				0.028			
4-Fctr Alpha$_{t-36,t-1}$ Medium (> 20th Ptile & ≤ 80th	0.001				0.011			
4-Fctr Alpha$_{t-36,t-1}$ High (> 80th Ptile)	0.916				0.359			
Total Return$_{t-36,t-1}$ Low (≤ 20th Ptile)	0.053				0.349			
Total Return$_{t-36,t-1}$ Medium (> 20th Ptile & ≤ 80th	0.049				0.008			
Total Return$_{t-36,t-1}$ High (> 80th Ptile)	0.038				0.250			

Table 4 – Determinants of fund performance

The table presents the results from regressions of monthly fund performance on lagged fund characteristics. The sample includes all U.S. domestic equity retail funds over the period 1997 to 2009. The dependent variable is the fund's 1 month forward-looking 4-factor alpha (Fama and French (1993); Carhart (1997)) using factor loadings estimated over the prior 36 months of data (t-1 to t-36). The independent variables include an intercept, the natural log of fund TNA and fund family TNA, the annual expense ratio, turnover calculated as the minimum of fund purchases and sales divided by TNA, an indicator variable of whether the fund is broker-sold (1=Yes) and whether it is an index fund (1=Yes). The independent variables also include indicator variables related to the existence of an institutional twin. *$Twin_{t-1}$ (=Yes for Inst. 1st/Inception Before Jan 1996)* is an indicator variable equal to one if the institutional twin was created before the retail fund or the institutional fund was created before the sample start date of January 1996. *$Twin_{t-1}$ (=Yes for Retail 1st - Before Inst. Fund Created)* is an indicator variable equal to one if the retail fund was created before the institutional fund, and if the monthly retail fund return observation predates the creation of the institutional twin fund. *$Twin_{t-1}$ (=Yes for Retail 1st - After Inst. Fund Created)* is an indicator variable equal to one if the retail fund was created before the institutional fund, and if the monthly retail fund return observation is measured after the creation of the institutional twin fund. Funds without an institutional twin have a 0 for all three indicator variables. Specification 3 only includes retail mutual fund twin returns for which there is no potential bias in the corresponding institutional twin observation (see section 3.1. for details). The reported results include date and fund family fixed effects as indicated at the bottom of the table. Standard errors are clustered by fund and date. The p-value of a difference in coefficients test between the before and after versions of the Retail 1st indicator variables is in the bottom row.

Regression:	1		2		3	
Log(Fund TNA)$_{t-1}$	-0.0249	(-3.7)	-0.0438	(-4.6)	-0.0427	(-4.6)
Log(Family TNA)$_{t-1}$	0.0093	(2.4)	-0.0657	(-3.4)	-0.0590	(-3.0)
Expense Ratio$_{t-1}$	-0.0116	(-0.4)	0.0143	(0.4)	0.0119	(0.3)
Turnover$_{t-1}$	-0.0001	(-1.3)	0.0000	(-0.2)	0.0000	(-0.2)
Broker-Sold ID$_{t-1}$ (=Yes)	-0.0473	(-2.8)	-0.0410	(-2.1)	-0.0413	(-2.0)
Index Fund ID$_{t-1}$ (=Yes)	0.0405	(1.0)	0.0820	(2.3)	0.0880	(2.4)
Twin$_{t-1}$ (=Yes for Retail 1st - Before Inst. Fund Created)	0.0265	(1.1)	0.0321	(1.4)	0.0719	(1.6)
Twin$_{t-1}$ (=Yes for Retail 1st - After Inst. Fund Created)	0.0978	(5.3)	0.1198	(5.2)	0.1452	(3.9)
Twin$_{t-1}$ (=Yes for Inst. 1st/Inception Before Jan 1996)	0.1086	(4.9)	0.1296	(5.9)	0.1040	(4.3)
Observations	214,360		214,360		192,148	
R-Squared	7.32%		7.85%		7.52%	
Calendar Fixed Effects	Yes		Yes		Yes	
Fund Family Fixed Effects	No		Yes		Yes	
Backfill/Survivorship Bias Filter	No		No		Yes	
Diff. Test p-Value	0.007		0.001		0.086	

Table 5 –Heckman correction, selection and performance regressions

Table 5 shows results from the first and second stage of a treatment effects model. The sample includes all U.S. domestic equity retail funds over the period 1997 to 2009. Panel A of the table shows a probit regression of whether or not a retail mutual fund has an institutional twin in any given month on characteristics of the retail fund (selection equation). The probit regression includes the 4-factor alpha (Fama and French (1993); Carhart (1997)) measured over the previous 36 months, the previous year's annual fund flow as a percent of TNA, and the natural log of fund TNA and fund family TNA, the annual expense ratio, turnover calculated as the minimum of fund purchases and sales divided by TNA, an indicator variable of whether the fund is broker-sold (1=Yes) and whether it is an index fund (1=Yes), all measured in the previous month. In addition we include the natural log of the family's institutional TNA estimated by aggregating the TNA of all separate accounts and institutional mutual funds listed in the Morningstar database for that fund family, an indicator variable for whether or not the family has any institutional assets under management in the Morningstar database, and the annualized tracking error, defined as the root mean square of the monthly return difference between the fund and the 4-factor adjusted benchmark, calculated over the previous 36 months. Panel B presents the results from regressions of monthly fund performance on lagged fund characteristics. The dependent variable is the fund's 1 month forward-looking 4-factor alpha (Fama and French (1993); Carhart (1997)) using factor loadings estimated over the prior 36 months of data (t-1 to t-36). The independent variables include an intercept, the natural log of fund TNA and fund family TNA, the annual expense ratio, turnover calculated as the minimum of fund purchases and sales divided by TNA, an indicator variable of whether the fund is broker-sold (1=Yes) and whether it is an index fund (1=Yes). Lambda is the Heckman correction and is based on the inverse Mills ratio. The independent variables also include indicator variables related to the existence of an institutional twin. *$Twin_{t-1}$ (=Yes for Retail 1st - Before Inst. Fund Created)* is an indicator variable equal to one if the retail fund was created before the institutional fund, and if the monthly retail fund return observation predates the creation of the institutional twin fund. *$Twin_{t-1}$ (=Yes for Retail 1st - After Inst. Fund Created)* is an indicator variable equal to one if the retail fund was created before the institutional fund, and if the monthly retail fund return observation is measured after the creation of the institutional twin fund. Funds without an institutional twin have a 0 for both indicator variables. Specification 3 only includes retail mutual fund twin returns for which there is no potential bias in the corresponding institutional twin observation (see section 3.1. for details). The reported results include date and fund family fixed effects as indicated at the bottom of the table. Standard errors are clustered by fund and date. The p-value of a difference in coefficients test between the before and after versions of the Retail 1st indicator variables is in the bottom row.

Panel A. First stage probit regression

	Coef.	t-Stat
Intercept	-9.404	(-17.2)
4-Factor Alpha$_{t-36,t-1}$	0.164	(2.4)
Annual Pct Flow$_{t-12,t-1}$	-0.015	(-0.8)
Log(Fund TNA)$_{t-1}$	0.162	(5.2)
Log(Family TNA)$_{t-1}$	-0.026	(-1.1)
Log(Family Institutional TNA)$_{t-1}$	0.016	(0.8)
Family Institutional TNA ID$_{t-1}$ (=Yes)	-0.326	(-0.7)
Expense Ratio$_{t-1}$	-0.021	(-0.2)
Turnover$_{t-12,t-1}$ (%)	0.000	(0.9)
Tracking Error$_{t-36,t-1}$ (4-Factor Model)	-6.770	(-1.2)
Broker-Sold ID$_{t-1}$ (=Yes)	0.229	(2.2)
Index Fund ID$_{t-1}$ (=Yes)	-0.266	(-1.2)
Observations	191,993	
Pseudo R-Squared	10.28%	
Calendar Fixed Effects	Yes	
Inv. Obj. Fixed Effects	Yes	

Table 5 –Heckman correction, selection and performance regressions (continued)

Panel B. Second stage fund performance regression

Regression:	1		2		3	
Log(Fund TNA)$_{t-1}$	-0.0722	(-3.8)	-0.0775	(-3.6)	-0.0753	(-3.4)
Log(Family TNA)$_{t-1}$	0.0135	(3.2)	-0.0606	(-3.0)	-0.0593	(-2.9)
Expense Ratio$_{t-1}$	-0.0035	(-0.1)	0.0129	(0.3)	0.0111	(0.3)
Turnover$_{t-1}$	-0.0001	(-1.4)	0.0000	(-0.2)	0.0000	(-0.3)
Lambda$_{t-1}$	-0.3299	(-2.8)	-0.2244	(-1.7)	-0.2203	(-1.6)
Broker-Sold ID$_{t-1}$ (=Yes)	-0.1119	(-4.1)	-0.0894	(-3.0)	-0.0875	(-2.9)
Index Fund ID$_{t-1}$ (=Yes)	0.1363	(2.7)	0.1498	(3.0)	0.1477	(2.9)
Twin$_{t-1}$ (=Yes for Retail 1st - Before Inst. Fund Created)	0.0258	(1.1)	0.0399	(1.7)	0.0769	(1.7)
Twin$_{t-1}$ (=Yes for Retail 1st - After Inst. Fund Created)	0.1037	(5.5)	0.1335	(5.9)	0.1596	(4.3)
Observations	191,993		191,993		180,512	
R-Squared	7.30%		7.83%		7.62%	
Calendar Fixed Effects	Yes		Yes		Yes	
Fund Family Fixed Effects	No		Yes		Yes	
Backfill/Survivorship Bias Filter	No		No		Yes	
Diff. Test p-Value	0.003		0.001		0.051	

Table 6 – Propensity score matching

Panel A of the table shows the results of a probit regression of whether or not an institutional twin is created for a given retail mutual fund in the following year, based on characteristics of the retail fund from the previous year. The t subscripts for the variables refer to months relative to the start of the year in which the institutional twin fund is created (i.e. for a twin creation any month in 2008, t-1 would refer to one month before January of 2008, or more simply, December of 2007). Using the propensity scores from this model, we construct a control sample of the 10 funds from the same time period with the closest propensity scores to the treatment group. In addition to variables previously described in the paper, the probit model also includes the natural log of the family's institutional TNA estimated by aggregating the TNA of all separate accounts and institutional mutual funds listed in the Morningstar database for that fund family, an indicator variable for whether or not the family has any institutional assets under management in the Morningstar database, and the annualized tracking error, defined as the root mean square of the monthly return difference between the fund and the 4-factor adjusted benchmark, calculated over the previous 36 months. Panel B compares the sample statistics for the treatment group and the control group. In addition to the mean and median, the table gives the p-values from a t-test of the difference in means and a sign test of the difference in medians.

Panel A. Propensity score probit estimation

	Coef.	t-Stat
Intercept	-5.349	(-8.8)
4-Factor Alpha$_{t-36,t-1}$	-0.035	(-0.4)
Annual Pct Flow$_{t-12,t-1}$	0.070	(2.3)
Log(Fund TNA)$_{t-1}$	0.132	(5.1)
Log(Family TNA)$_{t-1}$	-0.003	(-0.1)
Log(Family Institutional TNA)$_{t-1}$	-0.002	(-0.1)
Family Institutional TNA ID$_{t-1}$ (=Yes)	0.194	(0.4)
Expense Ratio$_{t-1}$	0.113	(1.3)
Turnover$_{t-12,t-1}$ (%)	0.000	(0.1)
Tracking Error$_{t-36,t-1}$ (4-Factor Model)	-2.229	(-0.4)
Broker-Sold ID$_{t-1}$ (=Yes)	0.206	(2.3)
Index Fund ID$_{t-1}$ (=Yes)	-0.082	(-0.5)
Observations	13,603	
Pseudo R-Squared	6.68%	
Calendar Fixed Effects	Yes	
Inv. Obj. Fixed Effects	Yes	

Panel B. Matched sample comparison

	Treatment		Control		Diff. Test p-value	
	Mean	Median	Mean	Median	Mean	Median
4-Factor Alpha (Annualized)	-0.37%	-0.86%	-0.17%	-0.77%	0.433	0.581
Quarterly Net Flow (%)	10.7%	1.0%	8.9%	0.2%	0.163	0.221
Log(Fund TNA)	20.4	20.4	20.5	20.4	0.377	0.976
Log(Family TNA)	22.7	23.0	22.8	23.3	0.340	0.269
Log(Family Institutional TNA)	15.9	22.6	16.3	22.7	0.313	0.922
Family Has Institutional Funds (=Yes)	67.0%	-	69.3%	-	0.178	-
Expense Ratio (Annual)	1.32%	1.26%	1.29%	1.26%	0.102	0.220
Turnover (%)	86.2%	70.5%	84.5%	62.0%	0.594	0.015
Tracking Error (4-Factor Model)	5.22%	4.60%	5.30%	4.66%	0.441	0.324
Broker-Sold (=Yes)	78.6%	-	76.9%	-	0.275	-
Index Fund (=Yes)	3.6%	-	4.4%	-	0.318	-

Table 7 – Performance, propensity-score matched sample

Table 7 contains results of a propensity score matched sample analysis of the change in performance for retail funds before and after an institutional twin is added (treatment group). The sample consists of 98 retail funds for which a twin institutional fund is created during our sample period, and for which we have return data in the three years before and after the creation of the institutional twin. The table includes results for 4-factor alphas over the 36 months before and after the event. To identify an appropriate control group, we use propensity scores from the probit model described in Table 6, Panel A. Using the propensity scores from this model, we construct several control samples. The results we report in the first block are based on nearest neighbor matching using ten observations from the same time period with the closest propensity score to create a control-group. The second (third) block shows results using a tolerance level on the maximum propensity score distance or caliper length of 0.005 to avoid the risk of bad matches. We then create a control group for each treated fund with up to 10 (5) observations that are within the maximum distance. In the fourth and fifth blocks, we further reduce the maximum propensity score distance to 0.001 and again use up to the 10 and 5 nearest observations. The last block of results is from a sample in which observations that are potentially subject to both backfill and survivorship biased have been removed and are based on nearest neighbor matching using ten observations to create a control-group. For each specification, the average outperformance for both the treatment and control group, before and after the addition of the institutional twin for the treatment group, is given. The differences between treatment and control groups before and after the event are given in the bottom (Treatment-Control) row. The differences in the before and after estimates for the treatment and control group are given in the last column (After-Before). The intersection of the bottom row and the last column gives the "Diff-in-Diff" estimate for each variable. The asterisks denote statistical significance in the following manner: *** significant at 1%; ** significant at 5%; and * significant at 10%.

		Caliper Length	Max # of Matches	# of Twin Funds	Total Matches	Before	After	(After-Before)
4-Factor Alpha (Annualized)	Treatment (Inst. Twin)					-0.373% **	-0.038%	0.335%
	Control (Matched Fund)	None	10	98	980	-0.170%	-1.383% ***	-1.213% ***
	(Treatment-Control)					-0.203%	1.345% ***	1.548% ***
4-Factor Alpha (Annualized)	Treatment (Inst. Twin)					-0.413% *	0.011%	0.423%
	Control (Matched Fund)	0.005	10	98	971	-0.260%	-1.481% ***	-1.221% ***
	(Treatment-Control)					-0.153%	1.491% ***	1.644% ***
4-Factor Alpha (Annualized)	Treatment (Inst. Twin)					-0.408%	0.005%	0.414%
	Control (Matched Fund)	0.005	5	98	481	-0.463%	-1.389% ***	-0.926% **
	(Treatment-Control)					0.055%	1.394% ***	1.339% **
4-Factor Alpha (Annualized)	Treatment (Inst. Twin)					-0.498% **	0.009%	0.507% *
	Control (Matched Fund)	0.001	10	98	949	-0.265%	-1.456% ***	-1.191% ***
	(Treatment-Control)					-0.233%	1.464% ***	1.698% ***
4-Factor Alpha (Annualized)	Treatment (Inst. Twin)					-0.485%	0.025%	0.510%
	Control (Matched Fund)	0.001	5	98	474	-0.473%	-1.360% ***	-0.888% **
	(Treatment-Control)					-0.012%	1.385% ***	1.398% **
4-Factor Alpha (Backfill/Survivorship Bias Filtered)	Treatment (Inst. Twin)					-0.406%	0.138%	0.543% *
	Control (Matched Fund)	None	10	53	530	0.132%	-1.276% ***	-1.408% ***
	(Treatment-Control)					-0.537% *	1.414% ***	1.951% ***

Table 8 – Mechanisms, propensity-score matched sample

Table 8 contains results of a matched sample analysis of the change in expense ratios and other fund characteristics for retail funds before and after an institutional twin is added (treatment group). To identify the matched sample, we use propensity scores from the probit model described in Table 6, Panel A. Using the propensity scores from this model, we construct a control sample of the 10 funds from the same time period with the closest propensity scores to each fund from the treatment group. For each variable the average for both the treatment and control group, before and after the addition of the institutional twin for the treatment group is given. The differences between treatment and control groups before and after the event are given in the bottom (Treatment-Control) row. The differences in the before and after estimates for the treatment and control group are given in the last (After-Before) column. The intersection of the bottom row and the last column gives the "Diff-in-Diff" estimate for each variable. The asterisks denote statistical significance in the following manner: *** significant at 1%; ** significant at 5%; and * significant at 10%. The first row shows results for the net annual expense ratio from the funds' semi-annual SEC N-SAR filing. The second row analyzes the return gap of Kacperczyk, Sialm and Zheng (2008), a measure of the hidden benefits/costs to a fund, calculated as the difference between the actual fund return and the return inferred from fund holdings (less expenses). The third and fourth rows show the brokerage commission rate calculated as the total brokerage commissions paid (Q21 for the series) divided by the sum of the manager's total purchases and sales (the sum of Q71.A and Q71.B for all funds in the series from the N-SAR filing), and the percent of funds using soft dollar or commission bundled payments to pay for fund distribution (Q26.A). The fifth row shows results for the Active Share measure of Petajisto (2010) and Cremers and Petajisto (2009), which is a measure of the overlap between the fund's holdings and the closest related index. The sixth row shows results for the value-weighted average number of analyst earnings estimates, where the weight is the percent of the equity portfolio held in each stock and the analyst following data is from IBES. For active share, the average over the 3 years before and after the event is analyzed. For all other variables, the average is taken over the year before and after the event.

		Obs.	Before	After	(After-Before)
Net Expense Ratio (NSAR)	Treatment (Inst. Twin)	82	1.215% ***	1.162% ***	-0.053% ***
	Control (Matched Fund)	820	1.268% ***	1.271% ***	0.003%
	(Treatment-Control)		-0.053% **	-0.109% ***	-0.056% ***
Return Gap (Annualized)	Treatment (Inst. Twin)	105	0.321% ***	0.579% ***	0.259%
	Control (Matched Fund)	1,050	-0.073%	-0.366% ***	-0.293% *
	(Treatment-Control)		0.394% **	0.946% ***	0.552% ***
Percent of Funds Using Soft Dollars for Distribution	Treatment (Inst. Twin)	82	23.26% ***	17.44% ***	-5.81% ***
	Control (Matched Fund)	820	21.40% ***	19.19% ***	-2.21% ***
	(Treatment-Control)		1.86%	-1.74%	-3.60% ***
Brokerage Commission Rate	Treatment (Inst. Twin)	82	0.096% ***	0.097% ***	0.000%
	Control (Matched Fund)	820	0.107% ***	0.151% ***	0.045%
	(Treatment-Control)		-0.011% **	-0.055%	-0.044%
Active Share (36 Month)	Treatment (Inst. Twin)	85	76.7% ***	76.4% ***	-0.30%
	Control (Matched Fund)	850	76.2% ***	75.3% ***	-0.97% ***
	(Treatment-Control)		0.44%	1.11%	0.67% *
Avg. # Analysts Estimates per Holding	Treatment (Inst. Twin)	112	13.71 ***	13.62 ***	-0.088 **
	Control (Matched Fund)	1,120	12.93 ***	13.04 ***	0.113 ***
	(Treatment-Control)		0.78 ***	0.57 ***	-0.201 ***

Table 9 –Heckman correction, selection and placebo performance regressions

Table 9 shows results from the first and second stage of a treatment effects model. The sample includes all U.S. domestic equity institutional funds (institutional mutual funds and separate accounts) in the Morningstar database over the period 1997 to 2009. Panel A of the table shows a probit regression of whether or not an institutional fund has a retail twin in any given month on characteristics of the institutional fund (selection equation). The probit regression includes the 4-factor alpha (Fama and French (1993); Carhart (1997)) measured over the previous 36 months, the previous year's annual fund flow as a percent of TNA, and the natural log of fund TNA and fund family's institutional TNA, the annual expense ratio, turnover calculated as the minimum of fund purchases and sales divided by TNA, an indicator variable of whether the fund is passively indexed (1=Yes), all measured in the previous month. In addition we include the natural log of the family's retail TNA estimated by aggregating the TNA of all retail mutual funds listed in the Morningstar database for that fund family, an indicator variable for whether or not the family has any retail assets under management in the Morningstar database, and the annualized tracking error, defined as the root mean square of the monthly return difference between the fund and the 4-factor adjusted benchmark, calculated over the previous 36 months. Panel B presents the results from regressions of monthly fund performance on lagged fund characteristics. The dependent variable is the fund's 1 month, forward-looking, 4-factor alpha (Fama and French (1993); Carhart (1997)) using factor loadings estimated over the prior 36 months of data (t-1 to t-36). The independent variables include an intercept, the natural log of fund TNA and fund family TNA, the annual expense ratio, turnover calculated as the minimum of fund purchases and sales divided by TNA, an indicator variable of whether the fund is passively indexed (1=Yes). Lambda is the Heckman correction and is based on the inverse Mills ratio. The independent variables also include indicator variables related to the existence of an retail twin. $Twin_{t-1}$ (=Yes for Inst.1st - Before Retail Fund Created) is an indicator variable equal to one if the retail fund was created before the institutional fund, and if the monthly retail fund return observation predates the creation of the institutional twin fund. $Twin_{t-1}$ (=Yes for Retail 1st - After Inst. Fund Created) is an indicator variable equal to one if the retail fund was created before the institutional fund, and if the monthly retail fund return observation is measured after the creation of the institutional twin fund.. Funds without a retail twin have a 0 for both indicator variables. Specification 3 only includes institutional fund twin returns for which there is no potential survivorship or backfill bias (see section 3.1. for details). The reported results include date and fund family fixed effects as indicated at the bottom of the table. Standard errors are clustered by fund and date. The p-value of a difference in coefficients test between the before and after versions of the Inst. 1st indicator variables is in the bottom row.

Panel A. First stage probit regression

	Coef.	t-Stat
Intercept	-3.326	(-4.5)
4-Factor $\text{Alpha}_{t-36,t-1}$	0.135	(1.9)
Annual Pct $\text{Flow}_{t-12,t-1}$	-0.004	(-0.4)
Log(Fund $\text{TNA})_{t-1}$	0.066	(2.2)
Log(Family Institutional $\text{TNA})_{t-1}$	-0.006	(-0.5)
Log(Family $\text{TNA})_{t-1}$	-0.031	(-1.2)
Family Retail TNA ID_{t-1} (=Yes)	1.327	(2.2)
Expense Ratio_{t-1}	-0.175	(-0.8)
$\text{Turnover}_{t-12,t-1}$ (%)	0.001	(1.0)
Tracking $\text{Error}_{t-36,t-1}$ (4-Factor Model)	-0.797	(-0.7)
Index Fund ID_{t-1} (=Yes)	0.272	(1.0)
Observations	98,185	
Pseudo R-Squared	7.32%	
Calendar Fixed Effects	Yes	
Inv. Obj. Fixed Effects	Yes	

Table 9 –Heckman correction, selection and placebo performance regressions (continued)

Panel B. Second Stage Fund Performance Regression

Regression:	1		2		3	
Log(Fund TNA)$_{t-1}$	-0.0248	(-3.4)	-0.0285	(-2.4)	-0.0292	(-3.1)
Log(Family Institutional TNA)$_{t-1}$	-0.0006	(-0.4)	-0.0235	(-2.8)	-0.0230	(-2.7)
Expense Ratio$_{t-1}$	-0.1540	(-4.1)	-0.0241	(-0.4)	-0.0413	(-0.7)
Turnover$_{t-1}$	-0.0005	(-1.3)	-0.0005	(-1.2)	-0.0007	(-2.1)
Lambda$_{t-1}$	0.0015	(0.0)	0.0504	(0.3)	0.0544	(0.4)
Index Fund ID$_{t-1}$ (=Yes)	0.0553	(1.3)	0.0407	(0.8)	0.0197	(0.5)
Twin$_{t-1}$ (=Yes for Inst. 1st - Before Retail Fund Created)	0.0271	(0.3)	-0.0297	(-0.3)	-0.0567	(-0.9)
Twin$_{t-1}$ (=Yes for Inst. 1st - After Retail Fund Created)	0.0190	(0.7)	0.0087	(0.3)	0.0102	(0.5)
Observations	98,185		98,185		82,271	
R-Squared	7.28%		8.30%		6.17%	
Calendar Fixed Effects	Yes		Yes		Yes	
Fund Family Fixed Effects	No		Yes		Yes	
Backfill/Survivorship Bias Filter	No		No		Yes	
Diff. Test p-Value	0.934		0.676		0.270	

Table 10 – Propensity score matching – placebo experiment

Panel A of the table shows the results of a probit regression of whether or not a retail twin is created for a given institutional fund in the following year, based on characteristics of the institutional fund from the previous year. The t subscripts for the variables refer to months relative to the start of the year in which the institutional twin fund is created (i.e. for a twin creation any month in 2008, t-1 would refer to one month before January of 2008, or more simply, December of 2007). Using the propensity scores from this model, we construct a control sample of the 10 institutional funds from the same time period with the closest propensity scores to the treatment group. In addition to variables previously described in the paper, the probit model also includes the natural log of the family's retail TNA estimated by aggregating the TNA of all retail mutual funds listed in the Morningstar database for that fund family, an indicator variable for whether or not the family has any retail assets under management in the Morningstar database, and the tracking error, defined as the root mean square of the monthly return difference between the fund and the 4-factor adjusted benchmark, calculated over the previous 36 months. Panel B compares the sample statistics for the treatment group and the control group. In addition to the mean and median, the table gives the p-values from a t-test of the difference in means and a sign test of the difference in medians.

Panel A. Propensity score probit estimation

	Coef.	t-Stat
Intercept	-2.487	(-3.6)
4-Factor Alpha$_{t-36,t-1}$	0.284	(3.2)
Quarterly Pct Flow$_{t-3,t-1}$	0.075	(1.4)
Log(Fund TNA)$_{t-1}$	-0.002	(-0.1)
Log(Family Institutional TNA)$_{t-1}$	-0.013	(-0.6)
Log(Family Retail TNA)$_{t-1}$	0.021	(4.3)
Family Retail TNA ID$_{t-1}$ (=Yes)	0.108	(0.2)
Expense Ratio$_{t-1}$	0.328	(1.9)
Tracking Error$_{t-36,t-1}$ (4-Factor Model)	4.410	(0.8)
Index Fund ID$_{t-1}$ (=Yes)	-0.103	(-0.5)
Observations	8,868	
Pseudo R-Squared	10.42%	
Calendar Fixed Effects	Yes	
Inv. Obj. Fixed Effects	Yes	

Panel B. Matched sample comparison

	Treatment		Control		Diff. Test p-value	
	Mean	Median	Mean	Median	Mean	Median
4-Factor Alpha (Annualized)	3.28%	1.51%	2.16%	0.72%	<0.001	0.003
Quarterly Net Flow (%)	19.6%	2.8%	17.2%	1.3%	0.511	0.430
Log(Fund TNA)	19.8	20.0	19.7	20.0	0.508	0.918
Log(Family Institutional TNA)	19.2	22.1	18.9	21.6	0.450	0.022
Log(Family Retail TNA)	15.5	21.1	15.5	20.2	0.952	0.510
Family Has Retail Funds (=Yes)	85.9%	-	85.1%	-	0.532	-
Expense Ratio	0.27%	0.08%	0.30%	0.07%	0.086	0.267
Tracking Error (4-Factor Model)	5.91%	4.83%	5.93%	4.85%	0.884	0.560
Index Fund (=Yes)	4.7%	-	6.0%	-	0.228	-

Table 11 – Performance analysis – placebo experiment

Table 11 contains results of a matched sample analysis of the change in performance and risk for institutional funds before and after a retail twin is added (treatment group). To identify the matched samples, we use propensity scores from the probit model described in Table 10. Using the propensity scores from this model, we construct a control sample of the 10 institutional funds from the same time period with the closest propensity scores to each fund from the treatment group. The last block of results is from a sample in which observations that are potentially subject to both backfill and survivorship biased have been removed and are based on nearest neighbor matching using ten observations to create a control-group. The results we report in the first block are based on nearest neighbor matching using ten observations from the same time period with the closest propensity score to create a control-group. The second block of results is from a sample in which observations that are potentially subject to both backfill and survivorship biased have been removed and are based on nearest neighbor matching using ten observations to create a control-group (see section 3.1 for details). The average for both the treatment and control group, before and after the addition of the retail twin for the treatment group is given. The differences between treatment and control groups before and after the event are given in the bottom (Treatment-Control) row. The differences in the before and after estimates for the treatment and control group are given in the last (After-Before) column. The intersection of this bottom row and the last column gives the "Diff-in-Diff" estimate for each variable. The asterisks denote statistical significance in the following manner: *** significant at 1%; ** significant at 5%; and * significant at 10%. The table includes results for 4-factor alphas over the 36 months before and after the event.

		Obs.	Before	After	(After-Before)
4-Factor Alpha (Annualized)	Treatment (Retail Twin)	85	3.28% ***	0.117%	-3.16% ***
	Control (Matched Fund)	850	2.16% ***	0.005%	-2.16% ***
	(Treatment-Control)		1.12% ***	0.112%	-1.01% ***
4-Factor Alpha (Backfill/Survivorship Bias Filtered)	Treatment (Retail Twin)	24	0.06%	-0.84% ***	-0.91% *
	Control (Matched Fund)	240	-0.12%	-0.59% *	-0.47%
	(Treatment-Control)		0.18%	-0.25%	-0.44%

Table 12 – Determinants of institutional share class flows

Table 12 presents coefficient estimates and t-statistics from pooled regressions of quarterly net fund flow on lagged fund characteristics. The regression is estimated using a sample that only includes retail and institutional funds from the matched twin sample over the period 1997 through 2009. Similar to specification 2 of Table 3, the sample includes retail funds with their corresponding separate account and institutional mutual funds twins. However, the sample also includes funds with retail and institutional share classes. For these funds, the retail share classes are aggregated and included as a separate retail fund-month observation and the institutional share classes are aggregated and included as a separate institutional fund-month observation. The dependent variable is the quarterly fund flow (t=0 to t=3) as a percent of TNA. The independent variables include an intercept, the natural log of lagged (t=-1) fund family and fund TNA, the lagged fund expense ratio, lagged turnover, the concurrent (t=0 to t=3) percentage quarterly flow to funds with the same investment objective (net of own fund flow), the lagged percentage quarterly fund flow, and two different measures of performance: 36-month total return and 4-factor alpha (Fama and French (1993); Carhart (1997)), calculated over the previous 36 months. For both performance measures, we allow for non-linearity through the use of a piece-wise linear performance specification with kinks at the 20th and 80th percentile of returns. For the total return measure, the percentiles are calculated within date and investment objective similar to Sirri and Tufano (1998), while the 4-factor alpha percentiles are calculated within date only. With these percentiles, the formula for the low return is LowRet=Min(0.2, RetPtile), the formula for the medium return is MedRet=Min(0.6, RetPtile-LowRet) , and the formula for the high return is HighRet=RetPtile-MedRet-LowRet. The regression specification allows for separately estimated coefficients for the three different fund types: retail funds, separate accounts/institutional mutual funds and the institutional share class funds. Difference in coefficient test p-values across the retail and the two different institutional fund type coefficients for the expense ratio, total return, and 4-factor alpha are provided at the bottom of the table. Standard errors are clustered by fund and by date and the total number of quarterly fund observations and the adjusted R-squared are provided.

Regression:	1					
	Retail Coef.		***SepAcct/InstMF***		***Inst ShrClss***	
Intercept	0.0908	(0.7)	0.4409	(0.7)	-0.0985	(-1.7)
$\text{Log(Family TNA)}_{t-1}$	0.0041	(8.2)	0.0001	(0.0)	0.0097	(7.7)
$\text{Log(Fund TNA)}_{t-1}$	-0.0118	(-15.1)	-0.0252	(-5.8)	-0.0282	(-14.5)
$\text{Expense Ratio}_{t-1}$	-0.0097	(-3.7)	-0.0646	(-3.1)	-0.0403	(-6.4)
Turnover_{t-1}	0.0002	(0.2)	-0.0002	(-0.1)	-0.0083	(-2.8)
$\text{InvObj Qtrly Pct Flow}_{t,t+3}$	0.4623	(7.5)	0.7874	(2.8)	0.6334	(7.0)
$\text{Qtrly Pct Flow}_{t-4,t-1}$	0.4101	(26.2)	0.0160	(0.5)	0.1761	(9.4)
$\text{Total Return}_{t-36,t-1}$ Low (≤ 20th Ptile)	0.0708	(7.9)	-0.0384	(-0.6)	0.1513	(5.1)
$\text{Total Return}_{t-36,t-1}$ Medium (> 20th Ptile & ≤ 80th Ptile)	0.0008	(0.2)	-0.0140	(-0.4)	0.0141	(1.2)
$\text{Total Return}_{t-36,t-1}$ High (> 80th Ptile)	0.1254	(7.4)	-0.0334	(-0.3)	0.1932	(4.9)
$\text{4-Fctr Alpha}_{t-36,t-1}$ Low (≤ 20th Ptile)	0.0477	(5.4)	0.2469	(3.6)	0.0590	(2.0)
$\text{4-Fctr Alpha}_{t-36,t-1}$ Medium (> 20th Ptile & ≤ 80th Ptile)	0.0188	(2.8)	0.0714	(1.5)	0.0467	(2.2)
$\text{4-Fctr Alpha}_{t-36,t-1}$ High (> 80th Ptile)	0.0844	(4.7)	0.1621	(3.0)	0.1426	(4.3)
Observations			86,137			
Adj. R-Squared			14.91%			

Coef. Difference Test p-Values (Retail vs. Institutional)	Retail vs. SepAcct/InstMF	Retail vs. Inst ShrClss
$\text{Expense Ratio}_{t-1}$	0.012	<0.001
$\text{4-Fctr Alpha}_{t-36,t-1}$ Low (≤ 20th Ptile)	0.004	0.709
$\text{4-Fctr Alpha}_{t-36,t-1}$ Medium (> 20th Ptile & ≤ 80th Ptile)	0.255	0.198
$\text{4-Fctr Alpha}_{t-36,t-1}$ High (> 80th Ptile)	0.160	0.077
$\text{Total Return}_{t-36,t-1}$ Low (≤ 20th Ptile)	0.087	0.004
$\text{Total Return}_{t-36,t-1}$ Medium (> 20th Ptile & ≤ 80th Ptile)	0.687	0.258
$\text{Total Return}_{t-36,t-1}$ High (> 80th Ptile)	0.148	0.084

Table 13 – Determinants of fund performance and institutional share class twins

Table 13 presents the results from regressions of monthly fund performance on lagged fund characteristics. The sample includes all U.S. domestic equity retail funds over the period 1997 to 2009. The dependent variable is the fund's 1 month forward-looking 4-factor alpha (Fama and French (1993); Carhart (1997)) using factor loadings estimated over the prior 36 months of data (t-1 to t-36). The independent variables include an intercept, the natural log of fund TNA and fund family TNA, the annual expense ratio, turnover calculated as the minimum of fund purchases and sales divided by TNA, an indicator variable of whether the fund is broker-sold (1=Yes) and whether it is an index fund (1=Yes). For retail funds with institutional share classes, the dependent and independent variables are calculated by value-weighting the quantities for the retail share classes only. The independent variables also include indicator variables related to the existence of an institutional twin. These indicator variables are separately identified for retail funds with separate account and institutional mutual fund twins (Inst. Funds) and retail funds with institutional share classes (Inst. ShrClss). For retail funds with institutional share classes, the twin identifier is defined based on the creation date of the institutional share class in the same way as for retail funds with separate account or institutional mutual fund twins. For the two different types of institutional twins, *$Twin_{t-1}$ (=Yes for Inst. 1st/Inception Before Jan 1996)* is an indicator variable equal to one if the institutional twin was created before the retail fund or the institutional fund was created before the sample start date of January 1996. *$Twin_{t-1}$ (=Yes for Retail 1st - Before Inst. Fund Created)* is an indicator variable equal to one if the retail fund was created before the institutional fund, and if the monthly retail fund return observation predates the creation of the institutional twin fund. *$Twin_{t-1}$ (=Yes for Retail 1st - After Inst. Fund Created)* is an indicator variable equal to one if the retail fund was created before the institutional fund, and if the monthly retail fund return observation is measured after the creation of the institutional twin fund. Funds without an institutional twin have a 0 for all six indicator variables. The reported results include date and fund family fixed effects as indicated at the bottom of the table. Standard errors are clustered by fund and date. The p-value of a difference in coefficients test between the before and after versions of the Retail 1st indicator variables is in the bottom row.

	Regression:	1		2	
	$Log(Fund\ TNA)_{t-1}$	-0.0251	(-4.0)	-0.0440	(-5.0)
	$Log(Family\ TNA)_{t-1}$	0.0068	(1.9)	-0.0558	(-3.0)
	$Expense\ Ratio_{t-1}$	-0.0390	(-1.4)	-0.0274	(-0.7)
	$Turnover_{t-1}$	-0.0001	(-1.2)	0.0000	(-0.2)
	Broker-Sold ID_{t-1} (=Yes)	-0.0423	(-2.4)	-0.0323	(-1.6)
	Index Fund ID_{t-1} (=Yes)	0.0342	(0.9)	0.0684	(1.9)
Inst. ShrClss	$Twin_{t-1}$ (=Yes for Retail 1st - Before Inst. Fund Created)	0.0877	(4.0)	0.0783	(3.4)
	$Twin_{t-1}$ (=Yes for Retail 1st - After Inst. Fund Created)	0.0199	(1.1)	0.0434	(2.0)
	$Twin_{t-1}$ (=Yes for Inst. 1st/Inception Before Jan 1996)	0.0228	(1.4)	0.0557	(3.0)
Inst. Funds	$Twin_{t-1}$ (=Yes for Retail 1st - Before Inst. Fund Created)	0.0436	(1.6)	0.0635	(2.2)
	$Twin_{t-1}$ (=Yes for Retail 1st - After Inst. Fund Created)	0.1124	(5.1)	0.1478	(5.7)
	$Twin_{t-1}$ (=Yes for Inst. 1st/Inception Before Jan 1996)	0.1272	(5.1)	0.1664	(6.4)
	Observations	208,068		208,068	
	R-Squared	7.26%		7.77%	
	Calendar Fixed Effects	Yes		Yes	
	Fund Family Fixed Effects	No		Yes	
	Diff. Test p-Value (Inst. Shareclasses)	0.003		0.180	
	Diff. Test p-Value (Inst. Funds)	0.014		0.003	

37

THE JOURNAL OF FINANCE • VOL. LXV, NO. 5 • OCTOBER 2010

Luck versus Skill in the Cross-Section of Mutual Fund Returns

EUGENE F. FAMA and KENNETH R. FRENCH*

ABSTRACT

The aggregate portfolio of actively managed U.S. equity mutual funds is close to the market portfolio, but the high costs of active management show up intact as lower returns to investors. Bootstrap simulations suggest that few funds produce benchmark-adjusted expected returns sufficient to cover their costs. If we add back the costs in fund expense ratios, there is evidence of inferior and superior performance (nonzero true α) in the extreme tails of the cross-section of mutual fund α estimates.

THERE IS A CONSTRAINT on the returns to active investing that we call equilibrium accounting. In short (details later), suppose that when returns are measured before costs (fees and other expenses), passive investors get passive returns, that is, they have zero α (abnormal expected return) relative to passive benchmarks. This means active investment must also be a zero sum game—aggregate α is zero before costs. Thus, if some active investors have positive α before costs, it is dollar for dollar at the expense of other active investors. After costs, that is, in terms of net returns to investors, active investment must be a negative sum game. (Sharpe (1991) calls this the arithmetic of active management.)

We examine mutual fund performance from the perspective of equilibrium accounting. For example, at the aggregate level, if the value-weight (VW) portfolio of active funds has a positive α before costs, we can infer that the VW portfolio of active investments outside mutual funds has a negative α. In other words, active mutual funds win at the expense of active investments outside mutual funds. We find that, in fact, the VW portfolio of active funds that invest primarily in U.S. equities is close to the market portfolio, and estimated before expenses, its α relative to common benchmarks is close to zero. Since the VW portfolio of active funds produces α close to zero in gross (pre-expense) returns, α estimated on the net (post-expense) returns realized by investors is negative by about the amount of fund expenses.

The aggregate results imply that if there are active mutual funds with positive true α, they are balanced by active funds with negative α. We test for the

*Fama is at the Booth School of Business, University of Chicago, and French is at the Amos Tuck School of Business Administration, Dartmouth College. We are grateful for the comments of Juhani Linnainmaa, Sunil Wahal, Jerry Zimmerman, and seminar participants at the University of Chicago, the California Institute of Technology, UCLA, and the Meckling Symposium at the University of Rochester. Special thanks to John Cochrane and the journal Editor, Associate Editor, and referees.

existence of such funds. The challenge is to distinguish skill from luck. Given the multitude of funds, many have extreme returns by chance. A common approach to this problem is to test for persistence in fund returns, that is, whether past winners continue to produce high returns and losers continue to underperform (see, e.g., Grinblatt and Titman (1992), Carhart (1997)). Persistence tests have an important weakness. Because they rank funds on short-term past performance, there may be little evidence of persistence because the allocation of funds to winner and loser portfolios is largely based on noise.

We take a different tack. We use long histories of individual fund returns and bootstrap simulations of return histories to infer the existence of superior and inferior funds. Specifically, we compare the actual cross-section of fund α estimates to the results from 10,000 bootstrap simulations of the cross-section. The returns of the funds in a simulation run have the properties of actual fund returns, except we set true α to zero in the return population from which simulation samples are drawn. The simulations thus describe the distribution of α estimates when there is no abnormal performance in fund returns. Comparing the distribution of α estimates from the simulations to the cross-section of α estimates for actual fund returns allows us to draw inferences about the existence of skilled managers.

For fund investors the simulation results are disheartening. When α is estimated on net returns to investors, the cross-section of precision-adjusted α estimates, $t(\alpha)$, suggests that few active funds produce benchmark-adjusted expected returns that cover their costs. Thus, if many managers have sufficient skill to cover costs, they are hidden by the mass of managers with insufficient skill. On a practical level, our results on long-term performance say that true α in net returns to investors is negative for most if not all active funds, including funds with strongly positive α estimates for their entire histories.

Mutual funds look better when returns are measured gross, that is, before the costs included in expense ratios. Comparing the cross-section of $t(\alpha)$ estimates from gross fund returns to the average cross-section from the simulations suggests that there are inferior managers whose actions reduce expected returns, and there are superior managers who enhance expected returns. If we assume that the cross-section of true α has a normal distribution with mean zero and standard deviation σ, then σ around 1.25% per year seems to capture the tails of the cross-section of α estimates for our full sample of actively managed funds.

The estimate of the standard deviation of true α, 1.25% per year, does not imply much skill. It suggests, for example, that fewer than 16% of funds have α greater than 1.25% per year (about 0.10% per month), and only about 2.3% have α greater than 2.50% per year (about 0.21% per month)—before expenses.

The simulation tests have power. If the cross-section of true α for gross fund returns is normal with mean zero, the simulations strongly suggest that the standard deviation of true α is between 0.75% and 1.75% per year. Thus, the simulations rule out values of σ rather close to our estimate, 1.25%. The power traces to the fact that a large cross-section of funds produces precise estimates of the percentiles of $t(\alpha)$ under different assumptions about σ, the standard deviation of true α. This precision allows us to put σ in a rather narrow range.

Readers suggest that our results are consistent with the predictions of Berk and Green (2004). We outline their model in Section II, after the tests on mutual fund aggregates (Section I) and before the bootstrap simulations (Sections III and IV). Our results reject most of their predictions about mutual fund returns. Given the prominence of their model, our contrary evidence seems an important contribution. The paper closest to ours is Kosowski et al. (2006). They run bootstrap simulations that appear to produce stronger evidence of manager skill. We contrast their tests and ours in Section V, after presenting our results. Section VI concludes.

I. The Performance of Aggregate Portfolios of U.S. Equity Mutual Funds

Our mutual fund sample is from the CRSP (Center for Research in Security Prices) database. We include only funds that invest primarily in U.S. common stocks, and we combine, with value weights, different classes of the same fund into a single fund (see French (2008)). To focus better on the performance of active managers, we exclude index funds from all our tests. The CRSP data start in 1962, but we concentrate on the period after 1983. During the period 1962 to 1983 about 15% of the funds on CRSP report only annual returns, and the average annual equal-weight (EW) return for these funds is 5.29% lower than for funds that report monthly returns. As a result, the EW average return on all funds is a nontrivial 0.65% per year lower than the EW return of funds that report monthly returns. Thus, during 1962 to 1983 there is selection bias in tests like ours that use only funds that report monthly returns. After 1983, almost all funds report monthly returns. (Elton, Gruber, and Blake (2001) discuss CRSP data problems for the period before 1984.)

A. The Regression Framework

Our main benchmark for evaluating fund performance is the three-factor model of Fama and French (1993), but we also show results for Carhart's (1997) four-factor model. To measure performance, these models use two variants of the time-series regression

$$R_{it} - R_{ft} = a_i + b_i(R_{Mt} - R_{ft}) + s_i SMB_t + h_i HML_t + m_i MOM_t + e_{it}. \quad (1)$$

In this regression, R_{it} is the return on fund i for month t, R_{ft} is the risk-free rate (the 1-month U.S. Treasury bill rate), R_{Mt} is the market return (the return on a VW portfolio of NYSE, Amex, and NASDAQ stocks), SMB_t and HML_t are the size and value-growth returns of Fama and French (1993), MOM_t is our version of Carhart's (1997) momentum return, a_i is the average return left unexplained by the benchmark model (the estimate of α_i), and e_{it} is the regression residual. The full version of (1) is Carhart's four-factor model, and the regression without MOM_t is the Fama–French three-factor model. The construction of SMB_t and HML_t follows Fama and French (1993). The momentum return,

MOM_t, is defined like HML_t, except that we sort on prior return rather than the book-to-market equity ratio. (See Table I below.)

Regression (1) allows a more precise statement of the constraints of equilibrium accounting. The VW aggregate of the U.S. equity portfolios of all investors is the market portfolio. It has a market slope equal to 1.0 in (1), zero slopes on the other explanatory returns, and a zero intercept—before investment costs. This means that if the VW aggregate portfolio of passive investors also has a zero intercept before costs, the VW aggregate portfolio of active investors must have a zero intercept. Thus, positive and negative intercepts among active investors must balance out—before costs.

There is controversy about whether the average SMB_t, HML_t, and MOM_t returns are rewards for risk or the result of mispricing. For our purposes, there is no need to take a stance on this issue. We can simply interpret SMB_t, HML_t, and MOM_t as diversified passive benchmark returns that capture patterns in average returns during our sample period, whatever the source of the average returns. Abstracting from the variation in returns associated with $R_{Mt} - R_{ft}$, SMB_t, HML_t, and MOM_t then allows us to focus better on the effects of active management (stock picking), which should show up in the three-factor and four-factor intercepts.

From an investment perspective, the slopes on the explanatory returns in (1) describe a diversified portfolio of passive benchmarks (including the risk-free security) that replicates the exposures of the fund on the left to common factors in returns. The regression intercept then measures the average return provided by a fund in excess of the return on a comparable passive portfolio. We interpret a positive expected intercept (true α) as good performance, and a negative expected intercept signals bad performance.[1]

Table I shows summary statistics for the explanatory returns in (1) for January 1984 through September 2006 (henceforth 1984 to 2006), the period used in our tests. The momentum factor (MOM_t) has the highest average return, 0.79% per month ($t = 3.01$), but the average values of the monthly market premium ($R_{Mt} - R_{ft}$) and the value-growth return (HML_t) are also large, 0.64% ($t = 2.42$) and 0.40% ($t = 2.10$), respectively. The size return, SMB_t, has the smallest average value, 0.03% per month ($t = 0.13$).

B. Regression Results for EW and VW Portfolios of Active Funds

Table II shows estimates of regression (1) for the monthly returns of 1984 to 2006 on EW and VW portfolios of the funds in our sample. In the VW portfolio, funds are weighted by assets under management (AUM) at the beginning of

[1] Formal justification for this definition of good and bad performance is provided by Dybvig and Ross (1985). Given a risk-free security, their Theorem 5 implies that if the intercept in (1) is positive, there is a portfolio with positive weight on fund i and the portfolio of the explanatory portfolios on the right of (1) that has a higher Sharpe ratio than the portfolio of the explanatory portfolios. Similarly, if the intercept is negative, there is a portfolio with negative weight on fund i that has a higher Sharpe ratio than the portfolio of the explanatory portfolios.

Table I

Summary Statistics for Monthly Explanatory Returns for the Three-Factor and Four-Factor Models

R_M is the return on a value-weight market portfolio of NYSE, Amex, and NASDAQ stocks, and R_f is the 1-month Treasury bill rate. The construction of SMB_t and HML_t follows Fama and French (1993). At the end of June of each year k, we sort stocks into two size groups. Small includes NYSE, Amex, and NASDAQ stocks with June market capitalization below the NYSE median and Big includes stocks with market cap above the NYSE median. We also sort stocks into three book-to-market equity (*B/M*) groups, Growth (NYSE, Amex, and NASDAQ stocks in the bottom 30% of NYSE *B/M*), Neutral (middle 40% of NYSE *B/M*), and Value (top 30% of NYSE *B/M*). Book equity is for the fiscal year ending in calendar year $k-1$, and the market cap in *B/M* is for the end of December of $k-1$. The intersection of the (independent) size and *B/M* sorts produces six value-weight portfolios, refreshed at the end of June each year. The size return, SMB_t, is the simple average of the month t returns on the three Small stock portfolios minus the average of the returns on the three Big stock portfolios. The value-growth return, HML_t, is the simple average of the returns on the two Value portfolios minus the average of the returns on the two Growth portfolios. The momentum return, MOM_t, is defined like HML_t, except that we sort on prior return rather than *B/M* and the momentum sort is refreshed monthly rather than annually. At the end of each month $t-1$ we sort NYSE stocks on the average of the 11 months of returns to the end of month $t-2$. (Dropping the return for month $t-1$ is common in the momentum literature.) We use the 30th and 70th NYSE percentiles to assign NYSE, Amex, and NASDAQ stocks to Low, Medium, and High momentum groups. The intersection of the size sort for the most recent June and the independent momentum sort produces six value-weight portfolios, refreshed monthly. The momentum return, MOM_t, is the simple average of the month t returns on the two High momentum portfolios minus the average of the returns on the two Low momentum portfolios. The table shows the average monthly return, the standard deviation of monthly returns, and the t-statistic for the average monthly return. The period is January 1984 through September 2006.

	Average Return				Standard Deviation				t-statistic			
	R_M-R_f	*SMB*	*HML*	*MOM*	R_M-R_f	*SMB*	*HML*	*MOM*	R_M-R_f	*SMB*	*HML*	*MOM*
1984–2006	0.64	0.03	0.40	0.79	4.36	3.38	3.17	4.35	2.42	0.13	2.10	3.01

Table II

Intercepts and Slopes in Variants of Regression (1) for Equal-Weight (EW) and Value-Weight (VW) Portfolios of Actively Managed Mutual Funds

The table shows the annualized intercepts ($12 * a$) and t-statistics for the intercepts ($t(Coef)$) for the CAPM, three-factor, and four-factor versions of regression (1) estimated on equal-weight (EW) and value-weight (VW) net and gross returns on the portfolios of actively managed mutual funds in our sample. The table also shows the regression slopes (b, s, h, and m, for $R_M - R_f$, SMB, HML, and MOM, respectively), t-statistics for the slopes, and the regression R^2, all of which are the same to two decimals for gross and net returns. For the market slope, $t(Coef)$ tests whether b is different from 1.0. Net returns are those received by investors. Gross returns are net returns plus 1/12th of a fund's expense ratio for the year. When a fund's expense ratio for a year is missing, we assume it is the same as other actively managed funds with similar assets under management (AUM). The period is January 1984 through September 2006. On average there are 1,308 funds and their average AUM is \$648.0 million.

	$12 * a$						
	Net	Gross	b	s	h	m	R^2
EW Returns							
Coef	−1.11	0.18	1.01				0.96
t(Coef)	−1.80	0.31	1.12				
Coef	−0.93	0.36	0.98	0.18	−0.00		0.98
t(Coef)	−2.13	0.85	−1.78	16.09	−0.24		
Coef	−0.92	0.39	0.98	0.18	−0.00	−0.00	0.98
t(Coef)	−2.05	0.90	−1.78	16.01	−0.25	−0.14	
VW Returns							
Coef	−1.13	−0.18	0.99				0.99
t(Coef)	−3.03	−0.49	−2.10				
Coef	−0.81	0.13	0.96	0.07	−0.03		0.99
t(Coef)	−2.50	0.40	−5.42	7.96	−3.22		
Coef	−1.00	−0.05	0.97	0.07	−0.03	0.02	0.99
t(Coef)	−3.02	−0.15	−5.03	7.78	−3.03	2.60	

each month. The EW portfolio weights funds equally each month. The intercepts in (1) for EW fund returns tell us whether funds on average produce returns different from those implied by their exposures to common factors in returns, whereas VW returns tell us about the fate of aggregate wealth invested in funds. Table II shows estimates of (1) for fund returns measured gross and net of fund expenses. Net returns are those received by investors. Monthly gross returns are net returns plus 1/12th of a fund's expense ratio for the year.

The market slopes in Table II are close to 1.0, which is not surprising since our sample is funds that invest primarily in U.S. stocks. The HML_t and MOM_t slopes are close to zero. Thus, in aggregate, active funds show little exposure to the value-growth and momentum factors. The EW portfolio of funds produces a larger SMB_t slope (0.18) than the VW portfolio (0.07). We infer that smaller funds are more likely to invest in small stocks, but total dollars invested in active funds (captured by VW returns) show little tilt toward small stocks.

The intercepts in the estimates of (1) summarize the average performance of funds (EW returns) and the performance of aggregate wealth invested in funds (VW returns) relative to passive benchmarks. In terms of net returns to investors, performance is poor. The three-factor and four-factor (annualized) intercepts for EW and VW net returns are negative, ranging from −0.81% to −1.00% per year, with t-statistics from −2.05 to −3.02. These results are in line with previous work (e.g., Jensen (1968), Malkiel (1995), Gruber (1996)).

The intercepts in (1) for EW and VW net fund returns tell us whether on average active managers have sufficient skill to generate returns that cover the costs funds impose on investors. Gross returns come closer to testing whether managers have any skill. For EW gross fund returns, the three-factor and four-factor intercepts for 1984 to 2006 are positive, 0.36% and 0.39% per year, but they are only 0.85 and 0.90 standard errors from zero. The intercepts in (1) for VW gross returns are quite close to zero, 0.13% per year ($t = 0.40$) for the three-factor version of (1), and −0.05% per year ($t = -0.15$) for the four-factor model.

Table II also shows estimates of the CAPM version of (1), in which $R_{Mt} - R_{ft}$ is the only explanatory return. The annualized CAPM intercept for VW gross fund returns for 1984 to 2006, −0.18% per year ($t = -0.49$), is again close to zero and similar to the estimates for the three-factor and four-factor models. It is not surprising that the intercepts of the three models are so similar (−0.18%, 0.13%, and −0.05% per year) since VW fund returns produce slopes close to zero for the non-market explanatory returns in (1).

We can offer an equilibrium accounting perspective on the results in Table II. When we add back the costs in expense ratios, α estimates for VW gross fund returns are close to zero. Thus, before expenses, there is no evidence that total wealth invested in active funds gets any benefits or suffers any losses from active management. VW fund returns also show little exposure to the size, value, and momentum returns, and the market return alone explains 99% of the variance of the monthly VW fund return. Together these facts say that during 1984 to 2006, active mutual funds in aggregate hold a portfolio that, before expenses, mimics market portfolio returns. The return to investors, however, is reduced by the high expense ratios of active funds. These results echo equilibrium accounting, but for a subset of investment managers where the implications of equilibrium accounting for aggregate investor returns need not hold.

C. Measurement Issues in the Tests on Gross Returns

The benchmark explanatory returns in (1) are before all costs. This is appropriate in tests on net fund returns where the issue addressed is whether managers have sufficient skill to produce expected returns that cover their costs. Gross returns pose more difficult measurement issues.

The issue in the tests on gross fund returns is whether managers have skill that causes expected returns to differ from those of comparable passive benchmarks. For this purpose, one would like fund returns measured before all costs and non-return revenues. This would put funds on the same pure return basis as the benchmark explanatory returns, so the regressions could focus on manager skill. Our gross fund returns are before the costs in expense ratios (including management fees), but they are net of other costs, primarily trading costs, and they include the typically small revenues from securities lending.

We could attempt to add trading costs to our estimates of gross fund returns. Funds do not report trading costs, however, and estimates are subject to large errors. For example, trading costs are likely to vary across funds because of differences in style tilts, trading skill, and the extent to which a fund demands immediacy in trade execution. Trading costs also vary through time. Our view is that estimates of trading costs for individual funds, especially actively managed funds, are fraught with error and potential bias, and are likely to be misleading. We prefer to stay with our simple definition of gross returns (net returns plus the costs in expense ratios), with periodic qualifications to our inferences.

An alternative approach (suggested by a referee) is to put the passive benchmarks produced by combining the explanatory returns in (1) in the same units as the gross fund returns on the left of (1). This involves taking account of the costs not covered in expense ratios that would be borne by an efficiently managed passive benchmark with the same style tilts as the fund whose gross returns are to be explained. Appendix A discusses this approach in detail. The bottom line is that for efficiently managed passive funds, the costs missed in expense ratios are close to zero. Thus, adjusting the benchmarks produced by (1) for estimates of these costs is unnecessary.

This does not mean our tests on gross fund returns capture the pure effects of skill. Though it appears that all substantial costs incurred by efficiently managed passive funds are in their expense ratios, this is less likely to be true for actively managed funds. The typical active fund trades more than the typical passive fund, and active funds are likely to demand immediacy in trading that pushes up costs. Our tests on gross returns thus produce α estimates that capture skill, less whatever net costs (costs minus non-return revenues) are missed by expense ratios. Equivalently, the tests say that a fund's management has skill only if it is sufficient to cover the missing costs (primarily trading costs). This seems like a reasonable definition of skill since an efficiently managed passive fund can apparently avoid these costs. More important, this is the definition of skill we can accurately test, given the unavoidable absence of accurate trading cost estimates for active funds.

The fact that our gross fund returns are net of the costs missed in expense ratios, however, does affect the inferences about equilibrium accounting we can draw from the aggregate results in Table II. Since the α estimates for VW gross fund returns in Table II are close to zero, they suggest that in aggregate funds show sufficient skill to produce expected returns that cover some or all of the costs missed in expense ratios. If this is the correct inference (precision is an issue), equilibrium accounting then says that the costs recovered by funds are matched by equivalent losses on investments outside mutual funds.

II. Berk and Green (2004)

Readers contend that our results (Table II and below) are consistent with Berk and Green (2004). Their model is attractive theory, but our results reject most of its predictions about mutual fund returns.

In their world, a fund is endowed with a permanent α, before costs, but it faces costs that are an increasing convex function of AUM. Investors use returns to update estimates of α. A fund with a positive expected α before costs attracts inflows until AUM reaches the point where expected α, net of costs, is zero. Outflows drive out funds with negative expected α. In equilibrium, all active funds (and thus funds in aggregate) have positive expected α before costs and zero expected α net of costs.

Our evidence that the aggregate portfolio of mutual funds has negative α net of costs contradicts the predictions of Berk and Green (2004). The results below on the net returns of individual funds also reject their prediction that all active managers have zero α net of costs. In fact, our results say that for most if not all funds, true α in net returns is negative.

Finally, equilibrium accounting poses a theoretical problem for Berk and Green (2004). Their model focuses on rational investors who optimally choose among passive and active alternatives. In aggregate, their investors have positive α before costs and zero α after costs. Equilibrium accounting, however, says that in aggregate investors have zero α before costs and negative α after costs.

III. Bootstrap Simulations

Table II says that, on average, active mutual funds do not produce gross returns above (or below) those of passive benchmarks. This may just mean that managers with skill that allows them to outperform the benchmarks are balanced by inferior managers who underperform. We turn now to simulations that use individual fund returns to infer the existence of superior and inferior managers.

A. Setup

To lessen the effects of "incubation bias" (see below), we limit the tests to funds that reach 5 million 2006 dollars in AUM. Since the AUM minimum is in 2006 dollars, we include a fund in 1984, for example, if it has more than about \$2.5 million in AUM in 1984. Once a fund passes the AUM minimum, it is included in all subsequent tests, so this requirement does not create selection bias. We also show results for funds after they pass \$250 million and \$1 billion. Since we estimate benchmark regressions for each fund, we limit the tests to funds that have at least 8 months of returns after they pass an AUM bound, so there is a bit of survival bias. To avoid having lots of new funds with short return histories, we only use funds that appear on CRSP at least 5 years before the end of our sample period.

Fund management companies commonly provide seed money to new funds to develop a return history. Incubation bias arises because funds typically open to the public—and their pre-release returns are included in mutual fund databases—only if the returns turn out to be attractive. The \$5 million AUM bound for admission to the tests alleviates this bias since AUM is likely to be low during the pre-release period.

Evans (2010) suggests that incubation bias can be minimized by using returns only after funds receive a ticker symbol from NASDAQ, which typically means they are available to the public. Systematic data on ticker symbol start dates are available only after 1998. We have replicated our tests for 1999 to 2006 using CRSP start dates for new funds (as in our reported results) and then using NASDAQ ticker dates (from Evans). Switching to ticker dates has almost no effect on aggregate fund returns (as in Table II), and has only trivial effects on the cross-section of $t(\alpha)$ estimates for funds (as in Table III below). We conclude that incubation bias is probably unimportant in our results for 1984 to 2006.

Our goal is to draw inferences about the cross-section of true α for active funds, specifically, whether the cross-section of α estimates suggests a world where true α is zero for all funds or whether there is nonzero true α, especially in the tails of the cross-section of α estimates. We are interested in answering this question for 12 different cross-sections of α estimates—for gross and net returns, for the three-factor and four-factor benchmarks, and for the three AUM samples. Thus, we use regression (1) to estimate each fund's three-factor or four-factor α for gross or net returns for the part of 1984 to 2006 after the fund passes each AUM bound.

The tests for nonzero true α in actual fund returns use bootstrap simulations on returns that have the properties of fund returns, except that true α is set to zero for every fund. To set α to zero, we subtract a fund's α estimate from its monthly returns. For example, to compute three-factor benchmark-adjusted gross returns for a fund in the \$5 million group, we subtract its three-factor α estimated from monthly gross returns for the part of 1984 to 2006 that the fund is in the \$5 million group from the fund's monthly gross returns for that period. We calculate benchmark-adjusted returns for the three-factor and four-factor models, for gross and net returns, and for the three AUM bounds. The result is 12 populations of benchmark-adjusted (zero-α) returns. (CAPM simulation results are in Appendix B.)

A simulation run is a random sample (with replacement) of 273 months, drawn from the 273 calendar months of January 1984 to September 2006. For each of the 12 sets of benchmark-adjusted returns, we estimate, fund by fund, the relevant benchmark model on the simulation draw of months of adjusted returns, dropping funds that are in the simulation run for less than 8 months. Each run thus produces 12 cross-sections of α estimates using the same random sample of months from 12 populations of adjusted (zero-α) fund returns.

We do 10,000 simulation runs to produce 12 distributions of t-statistics, $t(\alpha)$, for a world in which true α is zero. We focus on $t(\alpha)$, rather than estimates of α, to control for differences in precision due to differences in residual variance and in the number of months funds are in a simulation run.

Note that setting true α equal to zero builds different assumptions about skill into the tests on gross and net fund returns. For net returns, setting true α to zero leads to a world where every manager has sufficient skill to generate expected returns that cover all costs. In contrast, setting true α to zero in gross returns implies a world where every fund manager has just enough skill to produce expected returns that cover the costs missed in expense ratios.

Our simulation approach has an important advantage. Because a simulation run is the same random sample of months for all funds, the simulations capture the cross-correlation of fund returns and its effects on the distribution of $t(\alpha)$ estimates. Since we jointly sample fund and explanatory returns, we also capture any correlated heteroskedasticity of the explanatory returns and disturbances of a benchmark model. We shall see that these details of our approach are important for inferences about true α in actual fund returns.

Defining a simulation run as the same random sample of months for all funds also has a cost. If a fund is not in the tests for the entire 1984 to 2006 period, it is likely to show up in a simulation run for more or less than the number of months it is in our sample. This is not serious. We focus on $t(\alpha)$, and the distribution of $t(\alpha)$ estimates depends on the number of months funds are in a simulation run through a degrees of freedom effect. The distributions of $t(\alpha)$ estimates for funds that are oversampled in a simulation run have more degrees of freedom (and thinner extreme tails) than the distributions of $t(\alpha)$ for the actual returns of the funds. Within a simulation run, however, oversampling of some funds should roughly offset undersampling of others, so a simulation run should produce a representative sample of $t(\alpha)$ estimates for simulated returns that have the properties of actual fund returns, except that true α is zero for every fund. Oversampling and undersampling of fund returns in a simulation run should also about balance out in the 10,000 runs used in our inferences.

A qualification of this conclusion is in order. In a simulation run, as in the tests on actual returns, we discard funds that have less than 8 months of returns. This means we end up with a bit more oversampling of fund returns. As a result, the distributions of $t(\alpha)$ estimates in the simulations tend to have more degrees of freedom (and thinner tails) than the estimates for actual fund returns. This means our tests are a bit biased toward finding false evidence of performance in the tails of $t(\alpha)$ estimates for actual fund returns.

There are two additional caveats. (i) Random sampling of months in a simulation run preserves the cross-correlation of fund returns, but we lose any effects of autocorrelation. The literature on autocorrelation of stock returns (e.g., Fama (1965)) suggests that this is a minor problem. (ii) Because we randomly sample months, we also lose any effects of variation through time in the regression slopes in (1). (The issues posed by time-varying slopes are discussed by Ferson and Schadt (1996).) Capturing time variation in the regression slopes poses thorny problems, and we leave this potentially important issue for future research.

To develop perspective on the simulations, we first compare, in qualitative terms, the percentiles of the cross-section of $t(\alpha)$ estimates from actual fund returns and the average values of the percentiles from the simulations. We then turn to likelihood statements about whether the cross-section of $t(\alpha)$ estimates for actual fund returns points to the existence of skill.

B. First Impressions

When we estimate a benchmark model on the returns of each fund in an AUM group, we get a cross-section of $t(\alpha)$ estimates that can be ordered into a cumulative distribution function (CDF) of $t(\alpha)$ estimates for actual fund returns. A simulation run for the same combination of benchmark model and AUM group also produces a cross-section of $t(\alpha)$ estimates and its CDF for a world in which true α is zero. In our initial examination of the simulations we compare (i) the values of $t(\alpha)$ at selected percentiles of the CDF of the $t(\alpha)$ estimates from actual fund returns and (ii) the averages across the 10,000 simulation runs of the $t(\alpha)$ estimates at the same percentiles. For example, the first percentile of three-factor $t(\alpha)$ estimates for the net returns of funds in the \$5 million AUM group is −3.87, versus an average first percentile of −2.50 from the 10,000 three-factor simulation runs for the net returns of funds in this group (Table III).

For each combination of gross or net returns, AUM group, and benchmark model, Table III shows the CDF of $t(\alpha)$ estimates for actual returns and the average of the 10,000 simulation CDFs. The average simulation CDFs are similar for gross and net returns and for the two benchmark models. This is not surprising since true α is always zero in the simulations. The dispersion of the average simulation CDFs decreases from lower to higher AUM groups. This is at least in part a degrees of freedom effect; on average, funds in lower AUM groups have shorter sample periods.

B.1. Net Returns

The Berk and Green (2004) prediction that most fund managers have sufficient skill to cover their costs fares poorly in Table III. The left tail percentiles of the $t(\alpha)$ estimates from actual net fund returns are far below the corresponding average values from the simulations. For example, the 10^{th} percentiles of the actual $t(\alpha)$ estimates, −2.34, −2.37, and −2.53 for the \$5 million, \$250 million, and \$1 billion groups, are much more extreme than the average estimates from the simulation, −1.32, −1.31, and −1.30. The right tails of the $t(\alpha)$ estimates also do not suggest widespread skill sufficient to cover costs. In the tests that use the three-factor model, the $t(\alpha)$ estimates from the actual net returns of funds in the \$5 million group are below the average values from the simulations for all percentiles below the 98^{th}. For the \$1 billion group, only the 99^{th} percentile of three-factor $t(\alpha)$ for actual net fund returns is above the average simulation 99^{th} percentile, and then only slightly. For the \$250 million group, the percentiles of three-factor $t(\alpha)$ for actual net fund returns are all below the averages from the simulations. Figure 1 shows the actual and average simulated CDFs for the \$5 million AUM group.

Table III

Percentiles of $t(\alpha)$ Estimates for Actual and Simulated Fund Returns: January 1984 to September 2006

The table shows values of $t(\alpha)$ at selected percentiles (Pct) of the distribution of $t(\alpha)$ estimates for actual (Act) net and gross fund returns. The table also shows the percent of the 10,000 simulation runs that produce lower values of $t(\alpha)$ at the selected percentiles than those observed for actual fund returns (% < Act). Sim is the average value of $t(\alpha)$ at the selected percentiles from the simulations. The period is January 1984 to September 2006 and results are shown for the three- and four-factor models for the \$5 million, \$250 million, and \$1 billion AUM fund groups. There are 3,156 funds in the \$5 million group, 1,422 in the \$250 million group, and 660 in the \$1 billion group.

	5 Million			250 Million			1 Billion		
Pct	Sim	Act	%<Act	Sim	Act	%<Act	Sim	Act	%<Act
3-Factor Net Returns									
1	−2.50	−3.87	0.08	−2.45	−3.87	0.10	−2.39	−4.39	0.01
2	−2.17	−3.42	0.06	−2.13	−3.38	0.13	−2.09	−3.55	0.09
3	−1.97	−3.15	0.07	−1.94	−3.15	0.12	−1.91	−3.36	0.07
4	−1.83	−2.99	0.06	−1.80	−3.04	0.10	−1.78	−3.16	0.07
5	−1.71	−2.84	0.08	−1.69	−2.91	0.10	−1.67	−2.99	0.10
10	−1.32	−2.34	0.05	−1.31	−2.37	0.10	−1.30	−2.53	0.08
20	−0.87	−1.74	0.03	−0.86	−1.87	0.04	−0.86	−1.98	0.03
30	−0.54	−1.27	0.06	−0.54	−1.41	0.06	−0.54	−1.59	0.02
40	−0.26	−0.92	0.05	−0.27	−1.03	0.07	−0.27	−1.19	0.02
50	−0.01	−0.62	0.04	−0.01	−0.71	0.06	−0.01	−0.82	0.03
60	0.25	−0.29	0.11	0.25	−0.39	0.19	0.24	−0.51	0.05
70	0.52	0.08	0.51	0.52	−0.08	0.25	0.52	−0.20	0.08
80	0.85	0.50	3.20	0.84	0.37	1.68	0.84	0.25	0.85
90	1.30	1.01	8.17	1.29	0.89	5.19	1.28	0.82	4.81
95	1.68	1.54	30.55	1.66	1.36	14.17	1.64	1.34	17.73
96	1.80	1.71	40.06	1.76	1.49	17.24	1.74	1.52	26.33
97	1.94	1.91	49.35	1.90	1.69	25.92	1.87	1.79	42.86
98	2.13	2.17	58.70	2.08	1.90	30.43	2.04	2.02	50.07
99	2.45	2.47	57.42	2.36	2.29	43.92	2.31	2.40	63.11
4-Factor Net Returns									
1	−2.55	−3.94	0.04	−2.47	−3.94	0.08	−2.40	−4.22	0.01
2	−2.20	−3.43	0.04	−2.14	−3.43	0.09	−2.09	−3.48	0.08
3	−2.00	−3.08	0.13	−1.95	−3.07	0.25	−1.91	−3.11	0.23
4	−1.85	−2.88	0.13	−1.80	−2.88	0.22	−1.77	−2.95	0.21
5	−1.73	−2.74	0.12	−1.69	−2.78	0.18	−1.66	−2.86	0.14
10	−1.33	−2.23	0.14	−1.30	−2.34	0.14	−1.29	−2.48	0.07
20	−0.86	−1.67	0.10	−0.85	−1.80	0.11	−0.84	−1.96	0.05
30	−0.53	−1.25	0.12	−0.52	−1.39	0.10	−0.52	−1.54	0.04
40	−0.25	−0.88	0.21	−0.25	−1.04	0.14	−0.25	−1.23	0.05
50	0.01	−0.60	0.18	0.01	−0.76	0.11	0.01	−0.87	0.07
60	0.26	−0.29	0.25	0.27	−0.42	0.29	0.26	−0.49	0.19
70	0.54	0.02	0.37	0.54	−0.13	0.24	0.54	−0.18	0.24
80	0.87	0.44	1.76	0.86	0.27	0.72	0.86	0.17	0.45
90	1.33	1.04	10.62	1.31	0.86	4.40	1.30	0.86	7.07
95	1.72	1.53	23.82	1.69	1.37	14.35	1.67	1.31	14.13
96	1.84	1.67	28.21	1.80	1.51	18.23	1.78	1.45	17.16
97	1.99	1.84	31.30	1.94	1.65	18.62	1.91	1.57	17.05
98	2.19	2.09	39.12	2.12	1.79	15.57	2.08	1.76	18.86
99	2.52	2.40	36.96	2.42	2.22	29.88	2.36	2.26	42.00

(continued)

Table III—*Continued*

	5 Million			250 Million			1 Billion		
Pct	Sim	Act	%<Act	Sim	Act	%<Act	Sim	Act	%<Act
3-Factor Gross Returns									
1	−2.49	−3.07	4.11	−2.45	−3.16	3.16	−2.39	−3.29	1.88
2	−2.17	−2.68	4.79	−2.13	−2.67	6.01	−2.09	−2.70	5.64
3	−1.97	−2.48	4.20	−1.94	−2.51	4.47	−1.91	−2.51	5.12
4	−1.83	−2.31	4.41	−1.80	−2.35	4.68	−1.78	−2.33	5.77
5	−1.71	−2.19	4.15	−1.69	−2.18	5.99	−1.67	−2.18	6.52
10	−1.32	−1.72	5.75	−1.31	−1.77	5.94	−1.30	−1.86	4.15
20	−0.87	−1.10	13.61	−0.86	−1.24	7.18	−0.86	−1.43	2.52
30	−0.54	−0.71	20.03	−0.54	−0.79	15.10	−0.54	−1.00	4.28
40	−0.26	−0.36	29.74	−0.27	−0.43	23.84	−0.27	−0.59	10.25
50	−0.01	−0.06	38.87	−0.01	−0.15	26.28	−0.01	−0.28	13.48
60	0.25	0.28	56.05	0.25	0.14	31.47	0.24	0.05	21.21
70	0.52	0.63	71.81	0.52	0.48	43.62	0.52	0.35	26.70
80	0.85	1.06	85.21	0.84	0.88	58.14	0.84	0.79	44.31
90	1.30	1.59	90.01	1.29	1.41	69.39	1.28	1.34	60.63
95	1.68	2.04	92.10	1.66	1.81	72.89	1.64	1.78	70.37
96	1.80	2.20	93.73	1.76	1.93	73.44	1.74	1.96	77.00
97	1.94	2.44	95.97	1.90	2.19	84.36	1.87	2.22	85.47
98	2.13	2.72	97.29	2.08	2.47	89.30	2.04	2.37	83.72
99	2.45	3.03	96.66	2.36	2.83	90.95	2.31	2.97	94.63
4-Factor Gross Returns									
1	−2.55	−3.06	5.49	−2.47	−3.02	6.72	−2.40	−3.34	1.67
2	−2.20	−2.71	4.99	−2.14	−2.63	7.84	−2.09	−2.48	14.14
3	−2.00	−2.46	5.46	−1.95	−2.43	7.33	−1.91	−2.40	8.43
4	−1.85	−2.27	6.39	−1.80	−2.33	5.73	−1.77	−2.25	8.66
5	−1.73	−2.11	7.71	−1.69	−2.12	8.62	−1.66	−2.11	9.52
10	−1.33	−1.62	12.27	−1.30	−1.71	8.63	−1.29	−1.85	4.69
20	−0.86	−1.09	16.23	−0.85	−1.19	11.13	−0.84	−1.34	5.29
30	−0.53	−0.65	28.46	−0.52	−0.75	19.76	−0.52	−0.92	8.75
40	−0.25	−0.33	35.43	−0.25	−0.45	22.31	−0.25	−0.57	12.54
50	0.01	−0.02	44.53	0.01	−0.16	26.29	0.01	−0.29	14.40
60	0.26	0.28	53.17	0.27	0.09	25.86	0.26	0.05	22.48
70	0.54	0.62	64.90	0.54	0.48	43.11	0.54	0.36	27.78
80	0.87	0.98	70.19	0.86	0.85	50.07	0.86	0.82	47.07
90	1.33	1.58	84.76	1.31	1.36	58.66	1.30	1.41	65.72
95	1.72	2.05	88.77	1.69	1.87	73.81	1.67	1.83	70.55
96	1.84	2.21	91.03	1.80	2.01	76.27	1.78	1.95	70.91
97	1.99	2.39	92.01	1.94	2.21	81.22	1.91	2.04	66.61
98	2.19	2.58	91.20	2.12	2.43	83.35	2.08	2.30	74.26
99	2.52	3.01	93.44	2.42	2.72	81.41	2.36	2.57	71.98

Evidence of skill sufficient to cover costs is even weaker with an adjustment for momentum exposure. In the tests that use the four-factor model, the percentiles of the $t(\alpha)$ estimates for actual net fund returns are always below the average values from the simulations. In other words, the averages of the

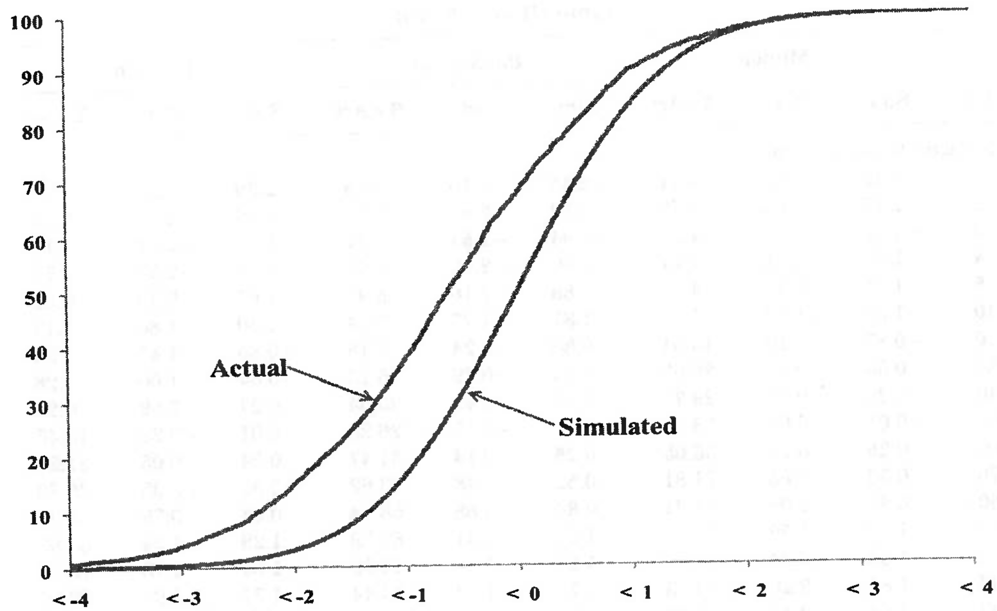


Figure 1. Simulated and actual cumulative density function of three-factor $t(\alpha)$ for net returns, 1984–2006.

percentile values of four-factor $t(\alpha)$ from the simulations of net returns (where by construction skill suffices to cover costs) always beat the corresponding percentiles of $t(\alpha)$ for actual net fund returns.

There is a glimmer of hope for investors in the tests on net returns. Even in the four-factor tests, the 99th and, for the \$5 million group, the 98th percentiles of the $t(\alpha)$ estimates for actual fund returns are close to the average values from the simulations. This suggests that some fund managers have enough skill to produce expected benchmark-adjusted net returns that cover costs. This is, however, a far cry from the prediction of Berk and Green (2004) that most if not all fund managers can cover their costs.

B.2. Gross Returns

It is possible that the fruits of skill do not show up more generally in net fund returns because they are absorbed by expenses. The tests on gross returns in Table III show that adding back the costs in expense ratios pushes up $t(\alpha)$ for actual fund returns. For all AUM groups, however, the left tail of three-factor $t(\alpha)$ estimates for actual gross fund returns is still to the left of the average from the simulations. For example, in the simulations the average value of the fifth percentile of $t(\alpha)$ for gross returns for the \$5 million group is −1.71, but the actual fifth percentile from actual fund returns is much lower, −2.19.

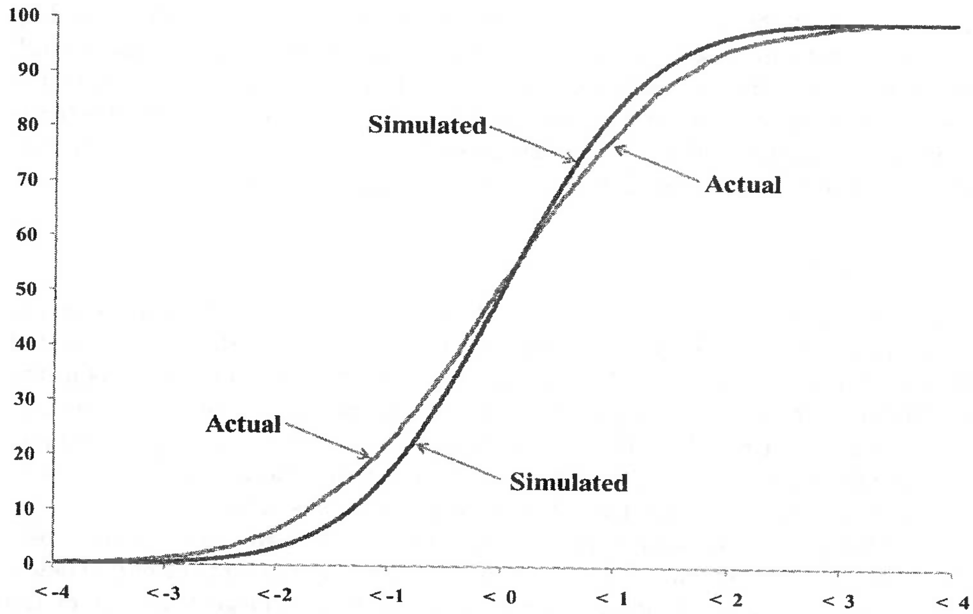


Figure 2. Simulated and actual cumulative density function of three-factor $t(\alpha)$ for gross returns, 1984–2006.

Thus, the left tails of the CDFs of three-factor $t(\alpha)$ suggest that when returns are measured before expenses, there are inferior fund managers whose actions result in negative true α relative to passive benchmarks.

Conversely, the right tails of three-factor $t(\alpha)$ suggest that there are superior managers who enhance expected returns relative to passive benchmarks. For the \$5 million AUM group, the CDF of $t(\alpha)$ estimates for actual gross fund returns moves to the right of the average from the simulations at about the 60^{th} percentile. For example, the 95^{th} percentile of $t(\alpha)$ for funds in the \$5 million group averages 1.68 in the simulations, but the actual 95^{th} percentile is higher, 2.04. For the two larger AUM groups the crossovers occur at higher percentiles, around the 80^{th} percentile for the \$250 million group and the 90^{th} percentile for the \$1 billion group. Figure 2 graphs the results for the three-factor benchmark and the \$5 million AUM group.

The four-factor results for gross returns in Table III are similar to the three-factor results, with a minor nuance. Adding a momentum control tends to shrink slightly the left and right tails of the cross-sections of $t(\alpha)$ estimates for actual fund returns. This suggests that funds with negative three-factor α estimates tend to have slight negative MOM_t exposure and funds with positive three-factor α tend to have slight positive exposure. Controlling for momentum pulls the α estimates toward zero, but only a bit.

Finally, the average simulation distribution of $t(\alpha)$ for the \$5 million fund group is like a t distribution with about 24 degrees of freedom. The average sample life of these funds is 112 months, so we can probably conclude that the simulation distributions of $t(\alpha)$ are more fat-tailed than can be explained by degrees of freedom. This may be due in part to fat-tailed distributions of stock returns (Fama (1965)). A referee suggests that active trading may also fatten the tails of fund returns. And properties of the joint distribution of fund returns may have important effects on the cross-section of $t(\alpha)$ estimates—a comment of some import in our later discussion of Kosowski et al. (2006).

C. Likelihoods

Comparing the percentiles of $t(\alpha)$ estimates for actual fund returns with the simulation averages gives hints about whether manager skill affects expected returns. Table III also provides likelihoods, in particular, the fractions of the 10,000 simulation runs that produce lower values of $t(\alpha)$ at selected percentiles than actual fund returns. These likelihoods allow us to judge more formally whether the tails of the cross-section of $t(\alpha)$ estimates for actual fund returns are extreme relative to what we observe when true α is zero.

Specifically, we infer that some managers lack skill sufficient to cover costs if low fractions of the simulation runs produce left tail percentiles of $t(\alpha)$ below those from actual net fund returns, or equivalently, if large fractions of the simulation runs beat the left tail $t(\alpha)$ estimates from actual net fund returns. Likewise, we infer that some managers produce benchmark-adjusted expected returns that more than cover costs if large fractions of the simulation runs produce right tail percentiles of $t(\alpha)$ below those from actual net fund returns. The logic is similar for gross returns, but the question is whether there are managers with skill sufficient to cover the costs (primarily trading costs) missing from expense ratios.

There are two problems in drawing inferences from the likelihoods in Table III. (i) Results are shown for many percentiles so there is a multiple comparisons issue. (ii) The likelihoods for different percentiles are correlated. One way to address these problems is to focus on a given percentile of each tail of $t(\alpha)$, for example, the 5^{th} and the 95^{th} percentiles, and draw inferences entirely from them. But this approach discards lots of information. We prefer to examine all the likelihoods, with emphasis on the extreme tails, where performance is most likely to be identified. As a result, our inferences from the formal likelihoods are somewhat informal.

C.1. Net Returns

The likelihoods in Table III confirm that skill sufficient to cover costs is rare. Below the 80^{th} percentile, the three-factor $t(\alpha)$ estimates for actual net fund returns beat those from the simulations in less than 1.0% of the net return simulation runs. For example, the 70^{th} percentile of the cross-section of three-factor $t(\alpha)$ estimates from the net returns of \$5 million funds (our full sample) is 0.08, and only 0.51% (about half of one percent) of the 10,000 simulation runs for this group produce 70^{th} percentile $t(\alpha)$ estimates below 0.08. It seems safe to conclude that most fund managers do not have enough skill to produce benchmark-adjusted net returns that cover costs. This again is bad news for Berk and Green (2004) since their model predicts that skill sufficient to cover costs is the general rule.

The likelihoods for the most extreme right tail percentiles of the three-factor $t(\alpha)$ estimates in Table III also confirm our earlier conclusion that some managers have sufficient skill to cover costs. For the \$5 million group, the 97th, 98th, and 99th percentiles of the cross-section of three-factor $t(\alpha)$ estimates from actual net fund returns are close to the average values from the simulations, and 49.35% to 58.70% of the $t(\alpha)$ estimates from the 10,000 simulation runs are below those from actual net returns. The likelihoods that the highest percentiles of the $t(\alpha)$ estimates from the net returns of funds in the \$5 million group beat those from the simulations drop below 40% when we use the four-factor model to measure α, but the likelihoods nevertheless suggest that some fund managers have enough skill to cover costs.

Some perspective is helpful. For the \$5 million group, about 30% of funds produce positive net return α estimates. The likelihoods in Table III tell us, however, that most of these funds are just lucky; their managers are not able to produce benchmark-adjusted expected returns that cover costs. For example, the 90th percentile of the $t(\alpha)$ estimates for actual net fund returns is near 1.00. The average standard error of the α estimates is 0.28% (monthly), which suggests that funds around the 90th percentile of $t(\alpha)$ beat our benchmarks by more than 3.3% per year for the entire period they are in the sample. These managers are sure to be anointed as highly skilled active investors. But about 90% of the net return simulation runs produce 90th percentiles of $t(\alpha)$ that beat those from actual fund returns. It thus seems that, like funds below the 90th percentile, most funds around the 90th percentile do not have managers with sufficient skill to cover costs; that is, true net return α is negative.

The odds that managers have enough skill to cover costs are better for funds at or above the 97th percentile of the $t(\alpha)$ estimates. In the \$5 million group, funds at the 97th, 98th, and 99th percentiles of three-factor $t(\alpha)$ estimates do about as well as would be expected if all fund managers were able to produce benchmark-adjusted expected returns that cover costs. But this just means that our estimate of true net return three-factor α for these funds is close to zero. If we switch to the four-factor model, the estimate of true α is negative for all percentiles of the $t(\alpha)$ estimates since the percentiles from actual net fund returns beat those from the simulations in less than 40% of the simulation runs.

What mix of active funds might generate the net return results in Table III? Suppose there are two groups of funds. Managers of good funds have just enough skill to produce zero α in net returns; bad funds have negative α. When the two groups are mixed, the expected cross-section of $t(\alpha)$ estimates is entirely to the left of the average of the cross-sections from the net return simulation runs (in which all managers have sufficient skill to cover costs). Even the extreme right tail of the $t(\alpha)$ estimates for actual net fund returns will be weighed down by bad managers who are extremely lucky but have smaller $t(\alpha)$ estimates than if they were extremely lucky good managers. In our tests, most of the cross-section of $t(\alpha)$ estimates for actual net fund returns is way left of what we expect if all managers have zero true α. Thus, most funds are probably in the negative true α group. At least for the \$5 million AUM sample, the 97^{th}, 98^{th}, and 99^{th} percentiles of the three-factor $t(\alpha)$ estimates for actual net fund returns are similar to the simulation averages. This suggests that buried in the results are fund managers with more than enough skill to cover costs, and the lucky among them pull up the extreme right tail of the net return $t(\alpha)$ estimates. Unfortunately, these good funds are indistinguishable from the lucky bad funds that land in the top percentiles of the $t(\alpha)$ estimates but have negative true α. As a result, our estimate of the three-factor net return α for a portfolio of the top three percentiles of the \$5 million group is near zero; the positive α of the lucky (but hidden) good funds is offset by the negative α of the lucky bad funds. And when we switch to the four-factor model, our estimate of true α turns negative even for the top three percentiles of the $t(\alpha)$ estimates.

Finally, our tests exclude index funds, but we can report that for 1984 to 2006 the net return three-factor α estimate for the VW portfolio of index funds (in which large, low cost funds get heavy weight) is -0.16% per year (-0.01% per month, $t = -0.61$), and four-factor α is 0.01% per year ($t = 0.02$). Since large, low cost index funds are not subject to the vagaries of active management, it seems reasonable to infer that the net return true α for a portfolio of these funds is close to zero. In other words, going forward we expect that a portfolio of low cost index funds will perform about as well as a portfolio of the top three percentiles of past active winners, and better than the rest of the active fund universe.

C.2. Gross Returns

The simulation tests for net returns ask whether active managers have sufficient skill to cover all their costs. In the tests on gross returns, the bar is lower. Specifically, the issue is whether managers have enough skill to at least cover the costs (primarily trading costs) missing from expense ratios.

The three-factor gross return simulations for the \$5 million AUM group suggest that most funds in the left tail of three-factor $t(\alpha)$ estimates do not have enough skill to produce benchmark-adjusted expected returns that cover trading costs, but many managers in the right tail have such skill. For the 40^{th} and lower percentiles, the three-factor $t(\alpha)$ estimates for the actual gross returns of funds in the \$5 million group beat those from the simulations in less than 30% of the simulation runs, falling to less than 6% for the 10^{th} and lower percentiles. Conversely, above the 60^{th} percentile, the three-factor $t(\alpha)$ estimates for actual gross fund returns beat those from the simulations in at least 56% of the simulation runs, rising to more than 90% for the 96^{th} and higher percentiles. As usual, the results are weaker when we switch from three-factor to four-factor benchmarks, but the general conclusions are the same.

For many readers, the important insight of Berk and Green (2004) is their assumption that there are diseconomies of scale in active management, not their detailed predictions about net fund returns (which are rejected in our tests).

The right tails of the $t(\alpha)$ estimates for gross returns suggest diseconomies. The extreme right tail percentiles of $t(\alpha)$ are typically lower for the \$250 million and \$1 billion groups than for the \$5 million group, and more of the simulation runs beat the extreme right tail percentiles of the $t(\alpha)$ estimates for the larger AUM funds. In the world of Berk and Green (2004), however, the weeding out of unskilled managers should also lead to left tails for $t(\alpha)$ estimates that are less extreme for larger funds. This prediction is not confirmed in our results. The left tails of the $t(\alpha)$ estimates for the \$250 million and \$1 billion groups are at least as extreme as the left tail for the \$5 million group. This contradiction in the left tails of the $t(\alpha)$ estimates makes us reluctant to interpret the right tails as evidence of diseconomies of scale.

The tests on gross returns point to the presence of skill (positive and negative). We next estimate the size of the skill effects. A side benefit is evidence on the power of the simulation tests.

IV. Estimating the Distribution of True α in Gross Fund Returns

To examine the likely size of the skill effects in gross fund returns we repeat the simulations but with α injected into fund returns. We then examine (i) how much α is necessary to reproduce the cross-section of $t(\alpha)$ estimates for actual gross fund returns, and (ii) levels of α too extreme to be consistent with the $t(\alpha)$ estimates for actual fund returns.

Given the evidence that, at least for the \$5 million group (our full sample), the distribution of $t(\alpha)$ estimates in gross fund returns is roughly symmetric about zero (Table III), it is reasonable to assume that true α is distributed around zero. It is also reasonable to assume that extreme levels of skill (good or bad) are rare. Concretely, we assume that each fund is endowed with a gross return α drawn from a normal distribution with a mean of zero and a standard deviation of σ per year.

The new simulations are much like the old. The first step again is to adjust the gross returns of each fund, setting α to zero for the three-factor and four-factor benchmarks and each of the three AUM groups. But now, before drawing the random sample of months for a simulation run, we draw a true α from a normal distribution with mean zero and standard deviation σ per year—the same α for every combination of benchmark model and AUM group for a given fund, but an independent drawing of α for each fund.

It seems reasonable that more diversified funds have less leeway to generate true α. To capture this idea, we scale the α drawn for a fund by the ratio of the fund's (three-factor or four-factor) residual standard error to the average standard error for all funds. We add the scaled α to the fund's benchmark-adjusted returns. We then draw a random sample (with replacement) of 273 months, and for each fund we estimate three-factor and four-factor regressions on the adjusted gross returns of the fund's three AUM samples. The simulations thus use returns that have the properties of actual fund returns, except we know true α has a normal distribution with mean zero and (for the "average" fund) standard deviation σ per year. We do 10,000 simulation runs, and a fund gets a new drawing of α in each run. To examine power, we vary σ, the standard deviation of true α, from 0.0% to 2.0% per year, in steps of 0.25%.

Table IV shows percentiles of the cross-section of $t(\alpha)$ estimates for actual gross fund returns (from Table III) and the average $t(\alpha)$ estimates at the same percentiles from the 10,000 simulation runs, for each value of σ. These are useful for judging how much dispersion in true α is consistent with the actual cross-section of $t(\alpha)$ estimates. For each σ, the table also shows the fraction of the simulation runs that produce percentiles of $t(\alpha)$ estimates below those from actual fund returns. We use these for inferences about the amount of dispersion in true α we might rule out as too extreme.

A. *Likely Levels of Performance*

If true α comes from a normal distribution with mean zero and standard deviation σ, Table IV provides two slightly different ways to infer the value of σ. We can look for the value of σ that produces average simulation percentile values of $t(\alpha)$ most like those from actual fund returns. Or we can look for the σ that produces simulation $t(\alpha)$ estimates below those for actual returns in about 50% of the simulation runs. If α has a normal distribution with mean zero and standard deviation σ, we expect the effects of the level of σ to become stronger as we look further into the tails of the cross-section of $t(\alpha)$. Thus, we are most interested in values of σ that match the extreme tails of the $t(\alpha)$ estimates for actual gross fund returns.

The normality assumption for true α is an approximation. We do not expect that a single value of σ (the standard deviation of true α) completely captures the tails of the $t(\alpha)$ estimates for actual fund returns, even if we allow a different σ for each tail. With this caveat, the three-factor and four-factor simulations for the \$5 million group suggest that σ around 1.25% to 1.50% per year captures the extreme left tail of the $t(\alpha)$ estimates for actual gross fund returns, and 1.25% works for the right tail. For the \$250 million and \$1 billion groups, the three-factor simulations again suggest σ around 1.25% to 1.50% per year for the left tail of the $t(\alpha)$ estimates for gross fund returns, but for the right tail σ is lower, 0.75% to 1.00% per year. In the four-factor simulations for the \$250 and \$1 billion groups $\sigma = 1.25\%$ per year seems to capture the extreme left tail of the $t(\alpha)$ estimates for gross fund returns, but the estimate of σ for the right tail is again lower, 0.75% per year. (To save space, Table IV shows results only for the \$5 million and \$1 billion AUM groups.)

The estimates do not suggest much performance, especially for larger funds. Thus, $\sigma = 1.25\%$ says that about one-sixth of funds have true gross return α greater than 1.25% per year (about 0.10% per month) and only about 2.4% have true α greater than 2.50% per year (0.21% per month). For perspective, the average of the OLS standard errors of individual fund α estimates—the average imprecision of α estimates—is 0.28% per month (3.4% per year). Moreover, much lower right tail σ estimates for the \$250 million and \$1 billion funds say that a lot of the right tail performance observed in the full (\$5 million) sample is due to tiny funds.

Table IV

Percentiles of $t(\alpha)$ Estimates for Actual and Simulated Gross Fund Returns with Injected α

The table shows values of $t(\alpha)$ at selected percentiles (Pct) of the distribution of $t(\alpha)$ estimates for Actual gross fund returns (repeated from Table III). The table also shows the average values of the $t(\alpha)$ estimates at the same percentiles from the 10,000 simulations, for seven values of σ (the annual standard deviation of injected α). The final seven columns of the table show, for each value of σ, the percent of the 10,000 simulation runs that produce lower $t(\alpha)$ estimates at the selected percentiles than actual fund returns. The period is January 1984 to September 2006 and results are shown for the three- and four-factor models for the $5 million and $1 billion AUM fund groups.

	Actual	Average $t(\alpha)$ from Simulations							Percent of Simulations below Actual						
Pct	$t(\alpha)$	0.50	0.75	1.00	1.25	1.50	1.75	2.00	0.50	0.75	1.00	1.25	1.50	1.75	2.00
3-Factor α, AUM > 5 Million															
1	−3.07	−2.63	−2.78	−2.99	−3.24	−3.54	−3.87	−4.23	7.46	15.74	36.37	69.29	92.13	99.00	99.94
2	−2.68	−2.27	−2.38	−2.52	−2.69	−2.89	−3.10	−3.34	8.03	13.67	26.30	49.41	76.25	93.77	99.04
3	−2.48	−2.06	−2.15	−2.27	−2.40	−2.55	−2.72	−2.91	6.55	10.94	19.32	35.16	57.73	80.85	94.54
4	−2.31	−1.91	−1.99	−2.08	−2.20	−2.33	−2.47	−2.62	6.85	10.63	17.71	30.94	49.77	71.63	88.96
5	−2.19	−1.78	−1.85	−1.94	−2.04	−2.16	−2.28	−2.41	6.36	9.68	15.54	26.25	41.95	61.44	80.67
10	−1.72	−1.37	−1.42	−1.48	−1.55	−1.62	−1.70	−1.78	7.86	10.82	15.39	22.37	32.27	44.75	59.63
90	1.59	1.35	1.40	1.46	1.53	1.60	1.68	1.76	86.35	81.64	74.23	64.06	51.23	36.75	22.61
95	2.04	1.75	1.83	1.92	2.02	2.13	2.26	2.39	88.27	82.46	72.10	56.14	36.57	18.02	5.63
96	2.20	1.87	1.96	2.06	2.17	2.30	2.45	2.60	90.76	85.40	74.75	57.87	36.04	16.06	4.08
97	2.44	2.03	2.12	2.23	2.37	2.53	2.70	2.88	93.73	89.76	80.72	63.35	38.59	15.13	3.39
98	2.72	2.23	2.35	2.49	2.66	2.85	3.07	3.31	95.29	91.75	82.56	61.96	32.18	8.75	1.24
99	3.03	2.58	2.74	2.95	3.20	3.49	3.82	4.18	93.48	85.84	63.90	29.57	5.82	0.41	0.02

	Actual	Average $t(\alpha)$ from Simulations							Percent of Simulations below Actual						
Pct	$t(\alpha)$	0.25	0.50	0.75	1.00	1.25	1.50	1.75	0.25	0.50	0.75	1.00	1.25	1.50	1.75
3-Factor α, AUM > 1 Billion															
1	−3.29	−2.42	−2.54	−2.73	−2.99	−3.31	−3.68	−4.09	2.20	3.63	8.89	22.71	48.69	76.60	92.24
2	−2.70	−2.12	−2.21	−2.34	−2.52	−2.73	−2.98	−3.25	6.27	9.11	15.83	28.78	51.16	75.26	90.50
3	−2.51	−1.94	−2.01	−2.12	−2.27	−2.44	−2.63	−2.84	5.82	8.10	13.21	22.97	39.66	61.21	80.93
4	−2.33	−1.80	−1.87	−1.97	−2.09	−2.23	−2.40	−2.57	6.43	8.75	13.73	22.41	36.16	55.28	74.51
5	−2.18	−1.69	−1.75	−1.84	−1.95	−2.08	−2.22	−2.37	7.42	9.81	14.45	22.65	35.12	51.86	70.08
10	−1.86	−1.32	−1.36	−1.42	−1.49	−1.58	−1.67	−1.77	4.46	5.54	7.88	11.48	17.22	25.15	36.41
90	1.34	1.29	1.34	1.40	1.47	1.55	1.64	1.74	58.48	52.67	43.86	34.00	23.16	13.78	6.93
95	1.78	1.66	1.72	1.81	1.92	2.04	2.18	2.33	67.79	60.84	49.40	35.18	20.70	9.71	3.14
96	1.96	1.77	1.83	1.93	2.05	2.19	2.35	2.53	74.69	68.02	56.45	40.71	24.10	11.07	3.42
97	2.22	1.90	1.97	2.08	2.23	2.39	2.58	2.78	84.12	79.12	68.67	52.24	32.34	14.98	4.56
98	2.37	2.07	2.16	2.29	2.46	2.67	2.90	3.16	82.02	75.12	61.59	41.62	21.00	6.82	1.56
99	2.97	2.35	2.46	2.64	2.88	3.18	3.52	3.90	93.92	90.90	81.49	61.18	32.94	11.99	2.84

(continued)

Table IV—*Continued*

Pct	Actual $t(\alpha)$	Average $t(\alpha)$ from Simulations 0.50	0.75	1.00	1.25	1.50	1.75	2.00	Percent of Simulations below Actual 0.50	0.75	1.00	1.25	1.50	1.75	2.00
4-Factor α, AUM > 5 Million															
1	−3.06	−2.69	−2.85	−3.06	−3.33	−3.63	−3.97	−4.34	10.99	22.26	47.07	78.71	95.82	99.64	99.97
2	−2.71	−2.31	−2.42	−2.57	−2.74	−2.94	−3.16	−3.41	8.36	14.61	28.65	52.04	78.54	94.44	99.25
3	−2.46	−2.09	−2.18	−2.30	−2.44	−2.60	−2.77	−2.96	8.82	14.08	25.32	42.58	66.01	86.35	96.53
4	−2.27	−1.93	−2.01	−2.11	−2.23	−2.36	−2.51	−2.66	9.85	15.06	24.90	39.73	60.10	80.07	93.38
5	−2.11	−1.80	−1.87	−1.96	−2.07	−2.18	−2.31	−2.45	11.46	16.87	26.22	39.83	57.77	76.75	90.48
10	−1.62	−1.38	−1.43	−1.49	−1.56	−1.64	−1.72	−1.80	16.05	21.02	27.97	37.70	49.46	63.12	76.66
90	1.58	1.38	1.43	1.50	1.56	1.64	1.72	1.80	79.81	74.21	66.46	55.96	43.49	30.25	18.17
95	2.05	1.80	1.87	1.96	2.07	2.18	2.31	2.44	83.71	76.60	66.13	50.01	31.99	15.33	4.86
96	2.21	1.92	2.00	2.11	2.22	2.36	2.50	2.66	86.46	79.25	68.28	50.91	30.64	12.88	3.36
97	2.39	2.08	2.17	2.29	2.43	2.59	2.76	2.95	87.10	79.52	66.70	46.44	24.15	7.38	1.23
98	2.58	2.29	2.41	2.55	2.72	2.92	3.14	3.38	85.21	75.29	57.57	32.34	10.71	1.74	0.15
99	3.01	2.66	2.81	3.02	3.28	3.58	3.91	4.27	88.10	75.85	49.98	19.30	3.25	0.19	0.01

Pct	Actual $t(\alpha)$	Average $t(\alpha)$ from Simulations 0.25	0.50	0.75	1.00	1.25	1.50	1.75	Percent of Simulations below Actual 0.25	0.50	0.75	1.00	1.25	1.50	1.75
4-Factor α, AUM > 1 Billion															
1	−3.34	−2.44	−2.56	−2.76	−3.03	−3.36	−3.74	−4.16	2.00	3.35	8.55	22.91	48.31	75.80	92.22
2	−2.48	−2.12	−2.22	−2.36	−2.54	−2.77	−3.02	−3.30	16.25	22.21	35.17	54.99	75.98	91.84	97.93
3	−2.40	−1.93	−2.01	−2.13	−2.28	−2.46	−2.66	−2.88	9.63	13.33	20.63	34.25	54.13	74.16	89.62
4	−2.25	−1.80	−1.87	−1.97	−2.10	−2.25	−2.42	−2.60	9.86	13.41	19.68	30.72	47.91	66.77	83.79
5	−2.11	−1.68	−1.75	−1.84	−1.96	−2.09	−2.24	−2.40	10.76	13.64	19.81	29.76	44.52	62.13	78.56
10	−1.85	−1.30	−1.35	−1.41	−1.49	−1.58	−1.67	−1.78	5.10	6.48	8.67	12.86	19.05	27.52	38.83
90	1.41	1.32	1.36	1.43	1.50	1.59	1.69	1.79	63.74	58.42	50.10	40.62	29.70	19.17	10.28
95	1.83	1.69	1.76	1.85	1.96	2.09	2.24	2.40	68.10	61.80	50.88	37.12	22.46	10.67	3.71
96	1.95	1.80	1.87	1.97	2.10	2.25	2.41	2.59	68.50	61.33	49.38	34.72	19.04	7.98	2.30
97	2.04	1.94	2.02	2.13	2.28	2.45	2.64	2.86	64.06	55.68	41.85	26.04	11.89	3.83	0.74
98	2.30	2.12	2.21	2.35	2.53	2.74	2.98	3.25	71.76	62.55	47.39	28.62	11.89	3.45	0.70
99	2.57	2.40	2.52	2.71	2.96	3.26	3.61	4.00	68.67	57.31	38.54	18.10	5.17	0.98	0.11

Our gross fund returns are net of trading costs. Returning trading costs to funds (if that is deemed appropriate) would increase the $t(\alpha)$ estimates in both the left and the right tails, which, depending on the (unknown) magnitudes, may move them toward more similar estimates of σ.

B. Unlikely Levels of Performance

What levels of σ can we reject? The answer depends on how confident we wish to be about our inferences. Suppose we are willing to accept a 20% chance of setting a lower bound for σ that is too high and a 20% chance of setting an upper bound that is too low. These bounds imply a narrower range than we would have with standard significance levels, but they are reasonable if our goal is to provide perspective on likely values of σ.

Under the 20% rule, the lower bound for the left tail estimate of σ is the value that produces left tail percentile $t(\alpha)$ estimates below those from actual fund returns in about 20% of the simulation runs. The upper bound for the left tail σ is the value that produces left tail percentiles of $t(\alpha)$ below those from actual fund returns in about 80% of the simulation runs. Conversely, under the 20% rule, the lower bound for the right tail σ estimate produces right tail percentile $t(\alpha)$ estimates below those from actual fund returns in about 80% of the simulation runs. And the upper bound for the right tail σ produces right tail percentiles of $t(\alpha)$ below those from actual fund returns in about 20% of the simulation runs.

In brief, applying the 20% rule leads to intervals for σ that are equal to the point estimates of the preceding section plus and minus 0.5%. For example, 1.25% per year works fairly well as the left tail estimate of σ for all AUM groups and for the three-factor and four-factor models, and the interval for the left tail σ estimates is 0.75% to 1.75%. For the \$5 million group, $\sigma = 1.25\%$ also works for the right tail, and the interval is again 0.75% to 1.75%. For the \$250 million and \$1 billion groups, the right tail estimate of σ drops to about 0.75% per year, and the 20% rule leads to an interval for σ from 0.25% to 1.25% per year.

What do these results say about the power of the simulation approach? The upper bound on σ for the \$5 million group, 1.75% per year, translates to a monthly σ for the cross-section of true α of about 0.146%. Suppose the standard error of each fund's α estimate is 0.28% per month (the sample average). With a monthly σ of 0.146%, the standard deviation of the cross-section of α estimates—caused by measurement error and dispersion in true α—is $(0.146^2 + 0.28^2)^{1/2} = 0.316\%$. This is only a bit bigger than 0.299%, the standard deviation implied by our estimate of σ for the \$5 million group, 1.25% per year. The fact that the simulations assign a relatively low probability to $\sigma \geq 1.75\%$ despite the small difference between the implied standard deviations of the α estimates for $\sigma = 1.25\%$ (the point estimate) and $\sigma = 1.75\%$ suggests that the simulations have power. The source of the power is our large sample of funds (3,156 in the \$5 million group). With so many funds, the percentiles of $t(\alpha)$

are estimated precisely, which produces power to draw inferences about σ. (We thank a referee for this insight.)

V. Kosowski et al. (2006)

The paper closest to ours is Kosowski et al. (2006). They use bootstrap simulations to draw inferences about performance in the cross-section of four-factor $t(\alpha)$ estimates for net fund returns. Their main inference is more positive than ours. They find that the 95$^{\text{th}}$ and higher percentiles of four-factor $t(\alpha)$ estimates for net fund returns are above the same simulation percentiles in more than 99% of simulation runs. This seems like strong evidence that among the best funds, many have more than sufficient skill to cover costs. Our simulations on net returns uncover much less evidence of skill. Two features of their tests account for their stronger results—simulation approach and time period.

We jointly sample fund (and explanatory) returns, whereas Kosowski et al. (2006) do independent simulations for each fund. The benefit of their approach is that the number of months a fund is in a simulation run always matches the fund's actual number of months of returns. The cost is that their simulations do not take account of the correlation of α estimates for different funds that arises because a benchmark model does not capture all common variation in fund returns. They summarize but do not show simulations that jointly sample the four-factor residuals of funds. But they never jointly sample fund returns and explanatory returns, which means (for example) they miss any effects of correlated movement in the volatilities of four-factor explanatory returns and residuals. In fact, in the results they show, the explanatory returns do not vary across simulation runs; the historical sequence of explanatory returns is used in every run.

Their rules for including funds in the simulation tests are also different. They include the complete return histories of all funds that survive more than 60 months (so there is survival bias). We include funds after they pass \$5 million in AUM if they have at least 8 months of returns thereafter (less survival bias).

Table V shows simulation results for their 1975 to 2002 period using (i) their rules for including funds and (ii) our rules. Note that both sets of simulations use our approach to drawing simulation samples, that is, a simulation run uses the same random sample of months for all funds, which allows for all effects implied by the joint distribution of fund returns, and of fund and explanatory returns.

The rules used to include funds affect the cross-section of $t(\alpha)$ estimates for actual fund returns. Specifically, the right tail $t(\alpha)$ estimates for actual fund returns are less extreme for our sample. This suggests that their rule that a fund must have at least 60 months of returns produces more survival bias than our 8-month rule. Another possibility is that some funds have high returns when they are tiny but do not do as well after they pass \$5 million. This may be due in part to an incubation bias in the fund sample of Kosowski et al. (2006),

Table V

Percentiles of Four-Factor $t(\alpha)$ for Actual and Simulated Fund Returns: 1975 to 2002

The table shows values of four-factor $t(\alpha)$ at selected percentiles (Pct) of the distribution of $t(\alpha)$ for actual (Act) net and gross fund returns for funds selected using the exclusion rules of Kosowski et al. (2006) and for funds in our $5 million AUM group selected using our exclusion rules. The period is 1975 to 2002 (as in Kosowski et al. (2006)). The table also shows the fraction (%<Act) of the 10,000 simulation runs that produce lower values of $t(\alpha)$ at the selected percentiles than those observed for actual fund returns. Sim is the average value of $t(\alpha)$ at the selected percentiles from the simulations.

	Kosowski et al. Exclusion Rules			Our Exclusion Rules		
Pct	Sim	Act	%<Act	Sim	Act	%<Act
1	−2.48	−3.69	0.18	−2.46	−3.70	0.16
2	−2.16	−3.25	0.19	−2.14	−3.17	0.30
3	−1.96	−2.87	0.53	−1.95	−2.80	0.70
4	−1.82	−2.55	1.34	−1.80	−2.63	0.69
5	−1.70	−2.36	1.90	−1.69	−2.41	1.36
10	−1.31	−1.92	2.17	−1.30	−1.95	1.66
20	−0.85	−1.41	2.15	−0.85	−1.41	2.17
30	−0.52	−1.01	3.18	−0.52	−1.00	3.54
40	−0.25	−0.65	5.75	−0.24	−0.66	5.35
50	0.01	−0.33	9.19	0.01	−0.34	8.50
60	0.27	−0.02	12.20	0.27	−0.03	11.92
70	0.55	0.29	16.51	0.55	0.27	14.86
80	0.87	0.73	32.80	0.87	0.69	28.11
90	1.32	1.44	68.19	1.32	1.34	56.29
95	1.69	1.97	82.42	1.69	1.81	68.32
96	1.80	2.18	88.38	1.80	2.00	75.70
97	1.94	2.38	90.73	1.94	2.25	83.74
98	2.12	2.59	91.38	2.12	2.51	87.57
99	2.40	3.07	95.79	2.42	2.83	88.37

since they include a fund's entire return history if the fund survives for 60 months.

For either sample of funds, joint sampling of fund returns (our approach) affects the simulation results. Kosowski et al. (2006) report that more than 99% of their simulation runs produce 95th percentile four-factor $t(\alpha)$ estimates below the 95th percentile from actual net fund returns. In Table V, the number drops to 82.42% for the fund sample selected using their rules and 68.32% using our rules. Skipping the details, we can report that the stronger performance results from the fund sample chosen using their rules is due to the 60-month survival rule. If the survival rule is reduced to 8 months, their rules for including funds produce simulation results close to ours. The important point, however, is that whatever inclusion rules are used, failure to account for the joint distribution of fund returns, and of fund and explanatory returns, biases the inferences of Kosowski et al. (2006) toward positive performance. (Cuthbertson, Nitzche, and O'Sullivan (2008) apply the simulation approach of Kosowski et al. to U.K. mutual funds, with similar results and, we guess, similar problems.)

Time period is also an important source of differences in results. Our simulations for 1984 to 2006 produce much less evidence of funds with sufficient skill to cover costs. In Table III, the CDFs of four-factor $t(\alpha)$ estimates for the net fund returns of 1984 to 2006 are always to the left of the average CDFs from the net return simulations (in which funds have sufficient skill to cover costs). Even in the extreme right tail of four-factor $t(\alpha)$ for net returns, more than 60% of the simulation runs beat the $t(\alpha)$ estimates for actual fund returns. But when our approach is applied to the 1975 to 2002 period of Kosowski et al. (2006), the 90^{th} and higher percentiles of $t(\alpha)$ for net fund returns are above the average values from the simulations (Table V). And for the 97^{th} and higher percentiles, less than 20% of the simulation runs beat the $t(\alpha)$ estimates for actual fund returns.

What do we make of the stronger results for 1975 to 2002 versus 1984 to 2006? One story is that in olden times there were fewer funds and a larger percentage of managers with sufficient skill to cover costs. Over time the skilled managers lost their edge or went on to more lucrative pursuits (e.g., hedge funds). Or perhaps, the entry of hordes of mediocre managers posing as skilled (Cremers and Petajisto (2009)) buries the tracks of true skill. Stronger results for 1975 to 2002 may also be due to biases in the CRSP data that are more prevalent in earlier years (Elton et al. (2001)). Whatever the explanation, the stronger evidence for performance during 1975 to 2002 is interesting, but irrelevant for today's investors.

VI. Conclusions

For 1984 to 2006, when the CRSP database is relatively free of biases, mutual fund investors in aggregate realize net returns that underperform CAPM, three-factor, and four-factor benchmarks by about the costs in expense ratios. Thus, if there are fund managers with enough skill to produce benchmark-adjusted expected returns that cover costs, their tracks are hidden in the aggregate results by the performance of managers with insufficient skill.

When we turn to individual funds, the challenge is to distinguish skill from luck. With 3,156 funds in our full ($5 million AUM) sample, some do extraordinarily well and some do extraordinarily poorly just by chance. To distinguish between luck and skill, we compare the distribution of $t(\alpha)$ estimates from actual fund returns with the distribution from bootstrap simulations in which all funds have zero true α. The tests on net returns say that few funds have enough skill to cover costs. The distribution of three-factor $t(\alpha)$ estimates from net fund returns is almost always to the left of the zero α distribution. The extreme right tail of the three-factor $t(\alpha)$ estimates for net fund returns, however, is roughly in line with the simulated distribution. This suggests that some managers do have sufficient skill to cover costs. But the estimate of net return three-factor true α is about zero even for the portfolio of funds in the top percentiles of historical three-factor $t(\alpha)$ estimates, and the estimate of four-factor true α is negative. Moreover, the estimate of true α for funds in the top percentiles is no better than the estimated α (also near zero) for large, efficiently managed passive funds.

The simulation results for gross fund returns say that when returns are measured before the costs in expense ratios, there is stronger evidence of manager skill, negative as well as positive. For our \$5 million AUM sample, true three-factor or four-factor gross return α seems to be symmetric about zero with a cross-section standard deviation of about 1.25% per year (about 10 basis points per month). For larger (\$250 million and \$1 billion AUM) funds, the standard deviation for the left tail is again about 1.25% per year, but the right tail standard deviation of true α falls to about 0.75%.

Appendix A: Measurement Issues in Gross Returns

The question in the tests on gross fund returns is whether managers have skill that causes expected returns to differ from those of comparable passive benchmarks. For this purpose, we would like to have fund returns measured before all costs but net of non-return income like revenues from securities lending. This would put funds on the same pure return basis as the benchmark explanatory returns, so the tests could focus on the effects of skill. Our gross fund returns are before the costs in expense ratios, but they are net of other costs, primarily trading costs, and they include income from securities lending.

We could attempt to add trading costs to our estimates of gross fund returns. Funds do not report trading costs, however, and even when turnover is available, estimates of trading costs are subject to large errors (Carhart (1997)). For example, trading costs are likely to vary across funds because of differences in style tilts, trading skill, and the extent to which a fund is actively managed and demands immediacy in trade execution. Trading costs can also vary through time because of changes in a fund's management and general changes in the costs of trading. All this leads us to conclude that estimates of trading costs for individual funds, especially actively managed funds, are fraught with error and potential bias, and so can be misleading. As a result, we do not take that route in our tests on gross returns.

An alternative approach (suggested by a referee) is to put the passive benchmarks produced by combining the explanatory returns in (1) in the same units as the gross fund returns on the left of (1). This involves taking account of the costs (primarily trading costs) not covered in expense ratios that would be borne by an efficiently managed passive benchmark with the same style tilts as the fund whose gross returns are to be explained.

Vanguard's index funds are good candidates for this exercise since, except for momentum, Vanguard provides index funds (Total Stock Market Index Fund, Growth Index Fund, Value Index Fund, Small-Cap Index Fund, Small-Cap Growth Index Fund, and Small-Cap Value Index Fund) that track well-defined target passive portfolios much like the market portfolio and the components of SMB_t and HML_t in (1). (We thank an Associate Editor for this insight.) Because the Vanguard index funds closely track their targets and stock picking skill is not an issue, we can estimate the average annual costs not included in a fund's expense ratio. Specifically, we add a fund's expense ratio to its reported average annual return for the 10 years through 2008 and then subtract the result from the average annual return of the fund's target for the same period. (The same calculation for an actively managed fund would include the effects of skill, as well as the costs not in expense ratios.) For every Vanguard index fund, this estimate of the costs missed in expense ratios is negative; that is, the fund's target return, which is before all costs, beats the fund's actual net return by less than the fund's expense ratio. If anything, Vanguard's small cap index funds do better on this score than its large cap funds—a clear warning that presumptions about trading costs can be misleading.

The Vanguard results are probably not unusual. We can report that the CAPM, three-factor, and four-factor α estimates for 1984 to 2006 for the net returns on a VW portfolio of index funds (which is dominated by large funds with low expense ratios) are close to zero, 0.08%, −0.16%, and 0.01% per year ($t = 0.18$, -0.61, and 0.02). In other words, in aggregate, wealth invested in index funds seems to earn average returns that cover costs, including trading costs.

Passive mutual funds that focus on momentum do not as yet exist, so we do not have estimates of trading costs for such funds. Existing work (Grundy and Martin (2001), Korajczyk and Sadka (2004)) suggests that the costs are significant. In our tests, however, the cross-sections of four-factor α estimates for funds are similar to the cross-sections of three-factor estimates, and the three-factor and four-factor tests produce much the same inferences. Given the large average MOM_t return, these results suggest that nontrivial long-term exposure to MOM_t is rare, so ignoring MOM_t trading costs is inconsequential. Moreover, the discussion of results in the text centers primarily on the three-factor model. The four-factor results are primarily a robustness check.

The Vanguard evidence and the results for a VW portfolio of index funds suggest that for the market and the components of SMB_t and HML_t, comparable efficiently managed passive mutual funds can enhance returns through trading, securities lending, and perhaps in other ways, so that their total costs are close to their expense ratios. Thus, our three-factor α estimates for the gross returns of funds would hardly change if we adjusted their passive benchmarks for the costs missed in expense ratios.

This does not mean our tests on gross returns capture the pure effects of skill. Though expense ratios seem to capture the total costs of efficiently managed passive funds, this is less likely to be true for actively managed funds. The typical active fund trades more than the typical passive fund, and active funds are likely to demand immediacy in trading that produces positive costs. Because of their high turnover, active funds also have fewer opportunities to generate revenues via securities lending (which are also trivial for the Vanguard funds). In short, it seems more likely that for active funds the costs not included in expense ratios are positive. Thus, our tests on the gross returns of funds produce α estimates that capture the effects of skill, less any costs missed by the expense ratios of the funds.

Equivalently, our tests on gross returns say that a fund's management has skill only if the fund's expected gross returns are sufficient to cover the costs (primarily trading costs) not included in its expense ratio. This is a reasonable definition of skill since a comparable efficiently managed passive fund would apparently avoid these costs. More important, this definition of skill is the only one we can accurately test in the absence of accurate estimates of the trading costs of active funds (impossible with available data).

It is fortuitous that efficiently managed passive benchmarks do not seem to have substantial costs missed in their expense ratios since accurate adjustment for such costs is nontrivial, perhaps impossible. For example, consider an actively managed small value fund. The passive benchmark for the fund produced by the three-factor version of (1) is likely to imply positive weights on the market, *SMB*, and *HML*, which implies positive weights on the market (M), small stocks (S), and value stocks (H) and negative weights on big stocks (B) and growth stocks (L). Suppose that (contrary to our estimates) efficiently managed passive funds have nontrivial trading costs. We might then increase the three-factor gross return α estimate for an active fund for the trading costs of the long positions in M, S, and H and the short positions in B and L that passively replicate the small value style of the active fund. But this is overkill. The three-factor model produces a passive clone for an actively managed fund by inefficiently combining five passive portfolios. A small value fund simply buys a diversified portfolio of small value stocks and only bears the trading costs of these stocks. As a result, even a passive small value fund evaluated with the three-factor model is likely to produce a positive α estimate if we enhance the estimate with positive trading costs for the five components of its three-factor clone.

If we wish to adjust the tests on gross returns for the trading costs of an efficiently managed passive fund with the same style tilts as the active fund to be evaluated, the correct procedure is to add an estimate of the trading costs of a comparable efficiently managed passive fund to the active fund's gross return α estimate. For example, a small value active fund would be reimbursed for the trading costs (more precisely, for all the costs missed in the expense ratio) of an efficiently managed passive fund with the same style tilts. This is nontrivial since a style group includes active funds with widely different style tilts, and we need an efficiently managed passive clone for every active fund. Fortunately, the costs missed in expense ratios are apparently close to zero for efficiently managed passive funds, and ignoring them (as we do in our tests) is inconsequential for inferences.

Appendix B: CAPM Bootstrap Simulations

Table AI replicates the bootstrap simulations in Table III for a CAPM benchmark, that is, regression (1) with the excess market return as the only explanatory variable. The CAPM results are different. The CAPM tests on net returns produce what seems like strong evidence that some fund managers have sufficient skill to cover costs. Thus, for percentiles above the 90$^{\text{th}}$, the CAPM $t(\alpha)$

Table AI
Percentiles of CAPM $t(\alpha)$ Estimates for Actual and Simulated Fund Returns

The table shows values of $t(\alpha)$ at selected percentiles (Pct) of the distribution of CAPM $t(\alpha)$ estimates for actual (Act) net and gross fund returns. The table also shows the percent of the 10,000 simulation runs that produce lower values of $t(\alpha)$ at the selected percentiles than those observed for actual fund returns (%<Act). Sim is the average value of $t(\alpha)$ at the selected percentiles from the simulations. The period is January 1984 to September 2006 and results are shown for the \$5 million, \$250 million, and \$1 billion AUM fund groups.

	5 Million			250 Million			1 Billion		
Pct	Sim	Act	%<Act	Sim	Act	%<Act	Sim	Act	%<Act
Net Returns									
1	−2.36	−3.72	0.25	−2.30	−3.70	0.40	−2.27	−4.10	0.08
2	−2.06	−3.28	0.45	−2.02	−3.29	0.58	−2.00	−3.50	0.24
3	−1.88	−3.00	0.64	−1.85	−3.02	0.79	−1.84	−3.29	0.23
4	−1.75	−2.84	0.62	−1.72	−2.92	0.65	−1.71	−3.18	0.19
5	−1.65	−2.69	0.74	−1.62	−2.76	0.77	−1.62	−3.00	0.27
10	−1.29	−2.16	1.08	−1.28	−2.18	1.64	−1.28	−2.47	0.46
20	−0.86	−1.48	1.93	−0.86	−1.58	1.98	−0.87	−1.79	0.70
30	−0.54	−1.05	2.09	−0.55	−1.11	2.30	−0.56	−1.35	0.44
40	−0.26	−0.65	3.84	−0.27	−0.75	2.50	−0.28	−0.88	0.48
50	0.00	−0.29	8.05	0.00	−0.36	5.23	−0.01	−0.46	1.29
60	0.26	0.08	20.79	0.26	0.06	19.86	0.26	−0.10	4.02
70	0.53	0.49	46.40	0.53	0.47	43.16	0.54	0.31	18.52
80	0.84	0.95	71.01	0.84	0.89	61.89	0.84	0.72	36.21
90	1.26	1.66	91.09	1.25	1.49	79.61	1.24	1.42	73.88
95	1.61	2.31	97.29	1.58	2.09	92.39	1.56	1.91	84.74
96	1.71	2.45	97.55	1.67	2.23	93.43	1.66	2.03	85.94
97	1.84	2.68	98.46	1.79	2.43	95.05	1.77	2.22	89.01
98	2.01	2.89	98.69	1.95	2.60	95.07	1.92	2.47	92.06
99	2.29	3.21	98.88	2.21	2.96	96.51	2.16	2.76	92.96
Gross Returns									
1	−2.36	−3.04	4.09	−2.30	−3.01	5.35	−2.27	−3.29	2.00
2	−2.06	−2.66	5.29	−2.02	−2.67	6.32	−2.00	−2.93	2.57
3	−1.88	−2.45	5.88	−1.85	−2.45	7.17	−1.84	−2.76	2.37
4	−1.75	−2.26	7.41	−1.72	−2.31	7.54	−1.71	−2.49	3.99
5	−1.65	−2.13	7.82	−1.62	−2.16	8.80	−1.62	−2.34	4.91
10	−1.29	−1.65	11.87	−1.28	−1.66	13.56	−1.28	−1.95	4.93
20	−0.86	−0.95	33.12	−0.86	−1.04	25.14	−0.87	−1.35	7.59
30	−0.54	−0.55	44.63	−0.55	−0.63	35.49	−0.56	−0.88	12.00
40	−0.26	−0.19	62.18	−0.27	−0.26	50.41	−0.28	−0.43	24.27
50	0.00	0.16	77.76	0.00	0.10	67.74	−0.01	−0.05	41.74
60	0.26	0.53	89.27	0.26	0.46	81.45	0.26	0.36	67.57
70	0.53	0.98	96.44	0.53	0.91	91.87	0.54	0.77	82.64
80	0.84	1.44	97.60	0.84	1.37	95.08	0.84	1.18	87.03
90	1.26	2.12	98.96	1.25	1.98	97.29	1.24	1.82	94.23
95	1.61	2.76	99.65	1.58	2.47	98.14	1.56	2.33	96.87
96	1.71	2.89	99.69	1.67	2.72	98.98	1.66	2.46	97.14
97	1.84	3.12	99.77	1.79	2.85	99.01	1.77	2.59	97.16
98	2.01	3.35	99.84	1.95	3.05	99.18	1.92	2.84	98.03
99	2.29	3.72	99.89	2.21	3.37	99.35	2.16	3.34	99.14

estimates for actual net fund returns are always above the averages from the net return simulations (in which all managers have sufficient skill to cover costs), and the $t(\alpha)$ estimates for actual fund returns typically beat those from the simulations in more than 80% of simulation runs. Relative to the three-factor and four-factor tests in Table III, the CAPM tests on gross returns in Table AI also produce what seems like stronger evidence that some managers have skill that leads to positive true α, while others have negative true α.

In fact, the CAPM results just illustrate well-known patterns in average returns that cause problems for the CAPM during our sample period. Actual mutual fund returns contain the effects of size, value-growth, and momentum tilts in fund portfolios that are missed by the CAPM. Thus, even passive funds that tilt toward small stocks, value stocks, or positive momentum stocks are likely to produce positive α estimates in CAPM tests, despite the fact that their managers make no effort to pick individual stocks. The CAPM simulations allow for the relation between average return and market exposure, but they wash out all other patterns in average returns when they subtract each fund's CAPM α estimate from its returns. As a result, the CAPM simulations say that actual fund returns have nonzero true α.

Which patterns in average returns left unexplained by the CAPM are most responsible for the differences between the CAPM simulation results and the results for the three-factor and four-factor models? Table III says that adding the momentum factor to the three-factor model has minor effects on estimates of $t(\alpha)$. Since the momentum return MOM_t has the highest average premium during our sample period, we infer that long-term exposure to momentum is probably rare among mutual funds. The average size (SMB_t) premium is trivial during our 1984 to 2006 sample period (0.03% per month, Table I), so size tilts probably are not driving the different results for the CAPM. That leaves the value (HML_t) premium as the focus of the story. Funds in the right tail of the CAPM $t(\alpha)$ estimates are more likely to have positive HML_t exposure that makes them look good in CAPM tests, and funds in the left tail are likely to have negative HML_t exposure.

In short, the CAPM tests are a lesson about how failure to account for common patterns in returns and average returns can affect inferences about the skill of fund managers.

REFERENCES

Berk, Jonathan B., and Richard C. Green, 2004, Mutual fund flows in rational markets, *Journal of Political Economy* 112, 1269–1295.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Cremers, Martijn, and Antti Petajisto, 2009, How active is your fund manager? A new measure that predicts performance, *Review of Financial Studies* 22, 3329–3365.

Dybvig, Philip H., and Stephen A. Ross, 1985, The analytics of performance measurement using a security market line, *Journal of Finance* 40, 401–416.

Elton, Edwin J., Martin J. Gruber, and Christopher R. Blake, 2001, A first look at the accuracy of the CRSP mutual fund database and a comparison of the CRSP and Morningstar mutual fund databases, *Journal of Finance* 56, 2415–2430.

Evans, Richard, 2010, Mutual fund incubatier, *Journal of Finance* 65, Forthcoming.

Fama, Eugene F., 1965, The behavior of stock market prices, *Journal of Business* 38, 34–105.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Ferson, Wayne E., and Rudi W. Schadt, 1996, Measuring fund strategy and performance in changing economic conditions, *Journal of Finance* 51, 425–462.

French, Kenneth R., 2008, The cost of active investing, *Journal of Finance* 63, 1537–1573.

Grinblatt, Mark, and Sheridan Titman, 1992, Performance persistence in mutual funds, *Journal of Finance* 47, 1977–1984.

Gruber, Martin J., 1996, Another puzzle: The growth of actively managed mutual funds, *Journal of Finance* 51, 783–810.

Grundy, Bruce D., and J. Spencer Martin, 2001, Understanding the nature of the risks and the sources of the rewards to momentum investing, *Journal of Financial Studies* 14, 29–78.

Jensen, Michael C., 1968, The performance of mutual funds in the period 1945–1964, *Journal of Finance* 23, 2033–2058.

Korajczyk, Robert A., and Ronnie Sadka, 2004, Are momentum profits robust to trading costs? *Journal of Finance* 59, 1039–1082.

Kosowski, Robert, Allan Timmermann, Russ Wermers, and Hal White, 2006, Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis, *Journal of Finance* 61, 2551–2595.

Malkiel, Burton G., 1995, Returns from investing in equity mutual funds: 1971–1991, *Journal of Finance* 50, 549–572.

O'Sullivan, N., 2008, UK mutual fund performance: Skill or luck? *Journal of Empirical Finance* 15, 613–634 [with K. Cuthbertson and D. Nitzsche].

Sharpe, William F., 1991, The arithmetic of active management, *Financial Analysts Journal* 47, 7–9.

38

HARVARD

JOHN M. OLIN CENTER FOR LAW, ECONOMICS, AND BUSINESS

THE LAW AND FINANCE OF BROKER-DEALER MARK-UPS

Allen Ferrell

April 6, 2011

Discussion Paper

Harvard Law School
Cambridge, MA 02138

This paper is also a discussion paper of the
John M. Olin Center's Program on Corporate Governance.

THE LAW AND FINANCE OF BROKER-DEALER MARK-UPS

ALLEN FERRELL*

Abstract

The prices charged retail customers by broker-dealers for less-liquid, lower-priced securities have been of long-standing regulatory concern. In particular, the National Association of Securities Dealers (succeeded now by the Financial Industry Regulatory Authority) has long had regulations prohibiting broker-dealers from charging excessive "mark-ups" and "mark-downs." This paper, using a unique dataset generously provided by the National Association of Securities Dealers tracking some 161,635 equity transactions involving fourteen broker-dealers and retail customers in largely less liquid, lower-priced securities over the course of the 2003-2005 period, provides the first comprehensive analysis of the determinants of the mark-ups and mark-downs charged by broker-dealers. In particular, the effect of broker-dealer solicitation, broker-dealer participation in the trade as a principal, stock price volatility, stock price level, trade volume and the bid-ask spread are examined on the size of mark-ups and mark-downs charged. This analysis is placed in the context of the law on mark-ups and mark-downs.

JEL Classification: G18, K22, K41

* Greenfield Professor of Securities Law, Harvard Law School. I would like to thank the generous support and assistance of the National Association of Securities Dealers (NASD). The research assistance of Rose Zhao is also gratefully acknowledged. I would also like to acknowledge the financial support of the John M. Olin Foundation in Law, Economics, and Business at Harvard Law School. This analysis was originally presented to the National Association of Securities Dealers in 2007.

I. INTRODUCTION

For many years, the National Association of Securities Dealers (NASD) (now the Financial Industry Regulatory Authority) as well as other regulators have long been concerned with excessive mark-ups and mark-downs being charged retail customers by broker-dealers. Indeed, the concern with excessive mark-ups and mark-downs has been a concern and a focus of regulatory attention since almost the very beginning of the NASD's existence as an organization. Particularly over the course of the last fifteen years, the NASD has placed a great deal of emphasis on broker-dealer sales practices and investor protection issues both in its examinations and investigations. Indeed, the NASD in recent years has levied substantial fines on leading investment banks for excessive mark-ups. The NASD, for example, fined in 2004 Goldman, Sachs & Co., Deutsche Bank Securities, Miller Tabak Roberts Securities, LLC, and Citigroup Global Markets Inc. each $5 million for, among other violations, charging excessive mark-ups and mark-downs to its customers.[1] In 2007, the Financial Industry Regulatory Authority (FINRA), which succeeded the NASD, announced that it fined Morgan Stanley & Co. $1.5 million and had ordered the firm to pay more than $4.6 million in restitution for rule violations relating to the sale of securities to retail customers at excessive prices.[2]

Despite the prominence of excessive mark-ups and mark-downs in broker-dealer regulation virtually no academic work has been done on the topic. The most comprehensive analysis of mark-ups and mark-downs was done by the REPORT OF SPECIAL STUDY OF SECURITIES MARKETS (hereinafter SPECIAL STUDY) back in 1964. Based on an analysis of a comprehensive sample of broker-dealer equity transactions (some 161,635 transactions in all) that includes information on mark-ups, mark-downs and commissions, generously provided to the author by the NASD while he was an academic fellow at the NASD, this paper will investigate the determinants of the size of equity mark-ups and mark-downs including the effects of broker-dealer solicitation and the price level of a security. It will also explore which price breakpoints best capture the

[1] See NASD Letters of Acceptance, Waiver, and Consent with Deutsche Bank Securities, Inc., Miller Tabak Roberts Securities LLC, Citigroup Global Markets Inc., and Goldman Sachs & Co. (July 28, 2004). Other violations included inadequate supervision and inadequate recordkeeping.

[2] FINRA, News Release (August 2, 2007).

effect of a security's price level on mark-ups and mark-downs and which factors are correlated with unusually large mark-ups and mark-downs. This analysis is important as which factors drive mark-ups and mark-downs, and to what extent, is a frequent subject of litigation and regulatory attention. This analysis was originally completed and presented to the NASD in 2007.

But before presenting the empirical results, the paper will first define what constitutes a "mark-up" and "mark-down" in Part II and then briefly summarize the law of mark-ups, focusing on NASD (now FINRA) regulations, in Part III. Part IV will then describe the unique dataset that will be utilized with descriptive statistics being presented in Part V. Finally Part VI will present the empirical analysis on the determinants of mark-ups and mark-downs with some concluding thoughts in Part VII.

II. THE DEFINITION OF "MARK-UP" AND "MARK-DOWN"

A "mark-up" is the difference between the price charged to a customer for a security by a broker-dealer acting as a principal (including riskless principal transactions) and that security's "prevailing market price."[3] Conversely, a "mark-down" is the difference between the price paid to a customer for a security by a broker-dealer, again acting as a principal, and the "prevailing market price" of the security. In contrast, a commission is the fee a broker-dealer charges for acting as an agent in ensuring execution of a customer's buy or sell order. As is obvious from these definitions, the concept of the "prevailing market price" plays a critical role in defining a "mark-up" or "mark-down," and, hence, the size of the mark-up or mark-down.

The "prevailing market price" typically means, as the SEC explained in the *Alstead, Dempsey* decision, the "price at which dealers' trade with one another, i.e., the current inter-dealer market."[4] In turn, the "best evidence" for the current inter-dealer market price is a "dealer's contemporaneous cost."[5] The most important caveat to this general statement concerning when to use a "dealer's contemporaneous cost" as the

---

[3] See *Press Chemical Investment Services Corporation*, 166 F.3d 529 (2d Cir. 1999)

[4] *Alstead, Dempsey & Co.*, Exchange Act Release No. 34-20825, 47 S.E.C. 1034, 1035 (April 5, 1984). If a market in a security is dominated and controlled by a broker-dealer the "prevailing market price" inquiry becomes more complicated. For a discussion, see NASD Notice to Members 92-16. Mark-ups/Mark-downs in Equity Securities.

[5] *Id.*

baseline is when a broker-dealer engaged in the sale or purchase of security is acting as a "market-maker."

A "market-maker," in turn, is defined in section 3(a)(38) of the Exchange Act of 1934 as "any dealer acting in the capacity of block positioner, and any dealer who with respect to a security, holds himself out (by entering quotations in an inter-dealer communications system or otherwise) as being willing to buy and sell such security for his own account on a regular or continuous basis." Much of the controversy over the scope of the definition of "market-maker" has revolved around the application of section 3(a)(38)'s language to broker-dealer activities in the bond market.[6] The source of much of this controversy is the fact that even broker-dealers that place capital at risk by maintaining an inventory of bonds to sell to customers do not necessarily disseminate firm bid and ask prices, whether through an inter-dealer communications system or otherwise, to market participants generally.

For "market-makers," so defined, the "prevailing market price," and hence the price off of which the size of the mark-up or mark-down will be computed, is the contemporaneous prices charged by market-makers in actual sales to (or purchases from) other dealers or, in the absence of such prices, validated quotations.[7] The need to validate quotations prior to their use as the "prevailing market price" was made clear at the very start of mark-up regulation back in the 1944 case of *Sherman Gleason & Co.*[8] Quotations can be validated by comparing the quotations for the period in question with actual inter-dealer transactions. If such inter-dealer transactions occur at "prices at or around the quoted offers" during a certain time period then the quotations will be considered validated for that time period.[9] So, for example, if the validated offer for a security was $10 then the "prevailing market price" for purposes of computing the prevailing market price for contemporaneous transactions would also be $10.

The rationale for not using a "dealer's contemporaneous cost" for market-makers as the "prevailing market price" was originally, and perhaps most clearly, laid out in 1964

---

[6] The SEC explained in greater detail the factors that are relevant to determining whether a broker-dealer is acting as a market-marker in *LSCO Securities Inc.*, SEC Exchange Act Release No. 26779 (May 3, 1989).
[7] The use of unvalidated quotations is "widely recognized as problematic." *Orkin*, 31 F.3d 1056, 1064.
[8] 15 S.E.C. 639 (rejecting the broker-dealer's argument that the prevailing market price can be determined by reference to unvalidated quotations reported in the newspaper).
[9] See *Peter Kisch*, 47 S.E.C. 802 (1982).

in the SPECIAL STUDY. The SPECIAL STUDY explained that if "the fairness of the mark-up is based on the [broker-dealer's] contemporaneous cost, the mark-up would be computed on the basis of its bid price. In the case of many low priced securities, where the spread between the inside bid and offer may be considerably more than 5 percent, the [broker-dealer] might then have to sell to is customers at a price less than its inside offer . . ."[10] This is due to the fact that in inter-dealer trades a broker-dealer will often purchase a security at its bid price (not its offer price), the price at which it advertises to other market participants the price at which it is willing to buy a security at, and the bid price will therefore constitute that broker-dealer's "contemporaneous cost." The use of the broker-dealer's bid price as the "prevailing market price" might result in inadequate compensation, so the reasoning goes, to broker-dealers for making a market in a security, and the liquidity this activity provides the market, particularly for less liquid securities where the spread (the difference between the bid and offer price) is likely to be substantial.[11]

III. BROKER-DEALER MARK-UP REGULATION

There are several different sources of regulation affecting broker-dealer equity mark-ups and mark-downs. Most importantly, for purposes of this paper, are the NASD regulations prohibiting excessive mark-ups and mark-downs. The disclosure requirements of Rule 10b-10 and the antifraud rule of Rule 10b-5 which also affect the size of mark-ups and mark-downs will also be briefly reviewed.

A. NASD Regulation

The SEC originally proposed (although ultimately withdrew the proposal) in 1942 that broker-dealers disclose to customers, at or before the completion of each transaction, the best possible price obtainable in the exercise of reasonable diligence. Interestingly, this disclosure obligation would have attached whether or not the transaction had been solicited by the broker-dealer. The NASD took a somewhat different approach electing to

---

[10] SPECIAL STUDY at 649.

[11] *Id.* at 611 ("The market maker . . . in addition to executing the transaction, provides marketability by assuming the risk of taking positions.").

directly prohibit the charging of excessive mark-ups and mark-downs.[12] To this end, the NASD in 1943 stated that it "shall be deemed conduct inconsistent with just and equitable principles of trade for a member to enter into any transaction with a customer at a price not reasonably related to the current market price of the security."[13] The NASD justified the imposition of this requirement based on the NASD's Rule of Fair Price Conduct Rule 2110 which states that NASD members shall "observe high standards of commercial honor and just and equitable principles of trade"[14] and NASD Conduct Rule 2440 which provides, in part, that "broker-dealers must buy or sell at prices that are fair . . ." Importantly, the NASD requirement that the prices broker-dealers charge customers be "reasonably related" to the security's market price applies not only to principal transactions, but also to trades in which the broker-dealer is merely acting as agent.[15] In other words, the prohibition against charging excessive prices includes commissions as well as what has been traditionally defined as "mark-ups" and "mark-downs." The prohibition on excessive prices is not automatically lifted even if the mark-up, mark-down or commission is disclosed to the customer.

But how does one determine whether a mark-up or mark-down (or commission) is excessive and therefore in violation of the NASD's rules? In 1943 the NASD conducted a membership-wide questionnaire examination relating to mark-ups in retail transactions. Approximately 82% of the NASD membership filled out the questionnaire. It found that 71% of transactions were made at mark-ups of 5% or less. Forty-seven percent of the transactions were made at mark-ups of 3% or less. Shortly after this survey, the NASD explained in a letter to its membership dated October 25, 1943 that the "District Business Conduct Committees have been instructed to enforce [the prohibition on excessive mark-ups] having in mind the percentage of profit on which 71% of the transactions [in the

---

[12] As will be discussed, there has been since 1977 SEC disclosure requirements relating to mark-ups and mark-downs with respect to equity transactions.

[13] Interpretation of the Board of Governors, NASD Mark-Up Policy, NASD Interpretation IM-2440. See also Exchange Act Release No. 24368, 52 Fed. Reg. 15575 at 15576 (April 21, 1987) (excessiveness turns on whether the price charged bears a "reasonable relationship to the prevailing market price.")

[14] Article III, Sec. 1.

[15] See, e.g., N.A.S.D. 7-15-50, E-132 ("the commission in an agency transaction must be fair"); id. at E-134 ("[T]he commission charged the customer must not be unfair . . ."); SPECIAL STUDY at 647 ("The policy applies both to sales from inventory and to riskless principal, and to agency as well as principal transactions")

survey] were effected."[16] In other words, mark-ups above 5% were more likely to run afoul of the mark-up rule, or at least require a greater justificatory basis, than mark-ups below this threshold. The NASD's position in its October 25, 1943 letter was approved by the SEC, including the so-called 5% guideline.[17]

However, the 5% guideline, rather than being the end of the analysis of whether a particular mark-up or mark-down is excessive, is just the start. As the NASD and courts have repeatedly emphasized, the 5% guideline is just that, a guideline rather than a hard and fast rule.[18] A determination of excessiveness turns on a "range of factors."[19] Consistent with the fact-specific nature of the inquiry, there have been a number of enforcement proceedings in which the mark-up and mark-down are less than 5% but have nevertheless been found to violate NASD rules. These factors, and the actual effect they have on mark-ups and mark-downs, will be explored in Part VI.

Excessive mark-ups and mark-downs trigger not only the NASD prohibition first outlined in its October 25, 1943 letter but may also constitute a violation of a broker-dealer's supervisory responsibilities. A broker-dealer has a statutory obligation pursuant to section 15(b)(4)(E) of the Exchange Act of 1934 to provide reasonable supervision with a view to avoiding violations of rules and regulations of their employees. This statutory obligation of supervision extends to ensuring that mark-ups being charged customers are not excessive.[20] Moreover, NASD Rule 3010 specifically requires member broker-dealers to supervise employees to ensure compliance with NASD rules. Of course, among these rules is the prohibition on excessive mark-ups and mark-downs.[21]

The NASD prohibition on excessive mark-ups is a useful complement to a Rule 10b-5 action and NASD Conduct Rule 2120, which contains the NASD's prohibition on fraud, as one can run afoul of the NASD prohibition on excessive mark-ups without having to establish that the wrongdoer acted with scienter. Establishing scienter for

---

[16] N.A.S.D. 7-15-50, E-126.
[17] See Exchange Act Release No. 3623.
[18] See, e.g., *Samuel B. Franklin & Co. v. SEC*, 290 F.2d 718, 725 (9th Cir. 1961) ("There is no hard and fast '5 percent rule'"); NASD Conduct Rule 2440 ("broker-dealers must buy or sell at prices that are fair, taking into account all relevant circumstances . . . "); Securities Exchange Act Release No. 3935 (rejecting argument that all mark-ups above 5% are necessarily excessive). Indeed, the NASD is prohibited by section 15A(b)(6) of the Exchange Act from setting schedules of prices or charges.
[19] *Press v. Chemical Inv. Servs. Corp.*, 988 F.Supp. 375, 385 (S.D.N.Y. 1997).
[20] See *In re W.J. Nolan & Co.*; Exchange Act Release No. 44,833, 75 S.E.C. Docket 1972 (Sept. 24, 2001).
[21] NASD Notice to Members 92-16, Mark-ups/Mark-downs in Equity Securities.

purposes of Rule 10b-5 and NASD Conduct Rule 2120 can be time-consuming and expensive.

B. Rule 10B-10 Disclosure Obligation

The SEC adopted in 1977 Rule 10b-10 which requires broker-dealers to provide certain disclosures on trade confirmations. Broker-dealers must disclose on the trade confirmation any "mark-up, mark-down or similar renumeration it receives" in riskless principal transactions involving equity securities when not acting as a market-maker. Importantly, there is no similar disclose requirement with respect to mark-ups and mark-downs in debt transactions. The SEC and courts have consistently taken the view that compliance with Rule 10b-10 is not a safe-harbor from Rule 10b-5 liability.[22]

C. Rule 10B-5 Liability

Since almost the beginning of the SEC's existence, the SEC has consistently taken the position that fraud occurs when the price being charged for a security by a broker-dealer bears no "reasonable relation" to the "prevailing price" in the absence of disclosure that would enable a customer to make an informed decision as to whether to enter a transaction with a broker-dealer charging a certain mark-up or mark-down.[23] The SEC's definition of fraud in the context of mark-ups was judicially endorsed several years later by the Second Circuit in *Charles Hughes & Co. v. SEC*, 139 F.2d 434 (1943) in which the court concluded that undisclosed mark-ups ranging from 16% to 41% "operated as a fraud and deceit upon the purchasers."[24] While the *Hughes* court noted that the broker-dealer in that case had a "special duty" to customers in light of, in part, the advice it had given customers that they should buy the stock, the SEC has long deemed undisclosed excessive mark-ups to be fraudulent even in the absence of any

---

[22] See *Grandon*, 147 F.3d 184.
[23] See *In re Duker v. Duker*, 8 S.E.C. 386, 388-89 (1939).
[24] Id. at 436.

solicitation.[25] Later courts, including the Second Circuit, have not limited the *Hughes* treatment of undisclosed excessive mark-ups as fraud to solicited transactions.[26]

In modern Rule 10b-5 jurisprudence, nondisclosure, even of a material fact, does not constitute fraud absent a duty to disclose.[27] Accordingly, courts have sometimes found that there exists an implied broker-dealer duty to disclose mark-ups when those mark-ups are excessive thereby ensuring nondisclosure will potentially result in Rule 10b-5 liability.[28] How this line of cases is impacted by the Supreme Court's decision in *Dura Pharmaceuticals v. Broudo*, 544 U.S. 336 (2005), and its emphasis on loss causation as an element of a Rule 10b-5 cause of action, is unclear.

IV. The Data

The NASD conducted a "sweep examination" of fourteen member broker-dealer firms. In these sweeps, the NASD collected from these firms a wide range information on all equity transactions handled by these broker-dealers over the course of a three month period. Specifically, the targeted broker-dealers were instructed to provide mark-up and mark-down information on every equity transaction they handled (whether as agent or principal) during this time period. In addition, the broker-dealers provided the NASD information concerning: (1) the number of shares being bought or sold; (2) the total value of the securities being bought or sold; (3) the price charged per share; (4) whether the broker-dealer was acting as a principal or agent; (5) any commissions charged; (6) the commission that should have been charged for the transaction according to the firm's own internal standard commission schedule; (7) whether the transaction was solicited by the broker-dealer; (8) the date of the transaction; (9) whether the transaction was a sale or a purchase; (10) whether the transaction occurred on the New York Stock Exchange or the American Stock Exchange; and (11) the name of the security being bought or sold.

The resulting dataset is unique in that it is the first time such comprehensive mark-up and mark-down data has ever been comprehensively gathered, by the NASD or

---

[25] See *United Sec. Corp.*, 15 S.E.C. 719, 727 (1944).

[26] See, e.g., *Grandon v. Merrill Lynch & Co.*, 147 F.3d 184, 189 (2nd Cir. 1998) (stating the holding of the *Hughes* opinion as "[t]here exists an implied representation that broker-dealers charge their customers securities prices that are reasonably related to the prices charged in an open and competitive market.")

[27] "Silence, absent a duty to disclose, is not misleading under Rule 10b-5." *Basic Inc. v. Levinson*, 485 U.S. 224, 239 n. 17 (1988).

[28] See *Grandon*, 147 F.3d at 193.

any other party, for a large set of transactions. In total, information on 161,635 equity transactions – 86,286 of which were transactions involving mark-ups and the other 75,349 of which were transactions involving mark-downs – was gathered from the fourteen broker-dealers.

It is important in interpreting this data to bear in mind that the fourteen broker-dealers were not randomly selected from the universe of NASD broker-dealer members. All fourteen firms share some important characteristics in common. The selected broker-dealers tended to be small- to medium-sized firms; had had some regulatory compliance issues of some sort in the past; and tended to deal primarily with retail investors selling or buying less liquid securities. As a result, the results presented in this paper cannot be generalized to broker-dealers across the board given the selected nature of the sample. The selected nature of the data can be seen in the fact that the median stock price of securities purchased by customers was only $5.50 (suggestive of these securities being less liquid), which is substantially lower than the median stock price for many securities, such as those traded on the NYSE or on NASDAQ's National Market System. Indeed, only 2% of the customer purchases in the sample were NYSE or AMEX transactions while only 1.5% of the customer sales were NYSE or AMEX transactions. At the same time, the data does cover a large number of transactions were mark-up regulation is likely to be most relevant: firms focusing on less liquid securities involving retail customers where the firms have faced regulatory compliance issues in the past.

The transactional information gathered by the NASD from the fourteen broker-dealer firms was merged with stock price information from the Center for Research in Securities Prices (CRSP). The transactions were merged with the CRSP database based on the security's CUSIP identifier and trade date. The information obtained from CRSP was (1) the security's trade volume for the day on which the transaction occurred; (2) the total number of shares outstanding of the security; (3) the bid-ask spread (in percentage terms) of the security; (4) the daily closing stock prices (both bid and ask) of the security for each trading day in the month in which the transaction in the security occurred.

The daily stock prices for the month in which the transaction occurred (item 4) was used to calculate a commonly used measure of stock price volatility. The volatility of the stock was measured as:

$$\text{Volatility} = \Sigma\,(p(i) - A)^2 / n$$

where A is the mean stock price for the month, p is the price of the stock (bid or ask) for a day *i*, and n is the number of trading days in the month. The volatility was measured using, alternatively, the ask and bid closing stock price. The results presented in Part VI were not qualitatively affected by the choice of whether to use the closing ask or bid price in calculating the stock's volatility.

After the data gathered by the NASD was merged with CRSP, the final database consisted of approximately 6 million observations.

V. Descriptive Statistics

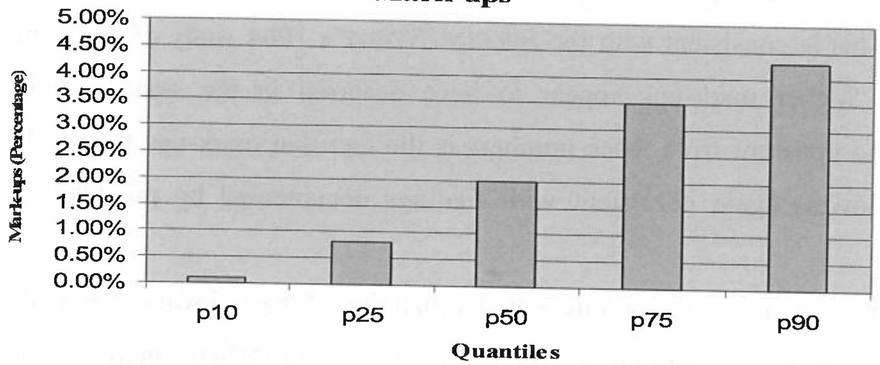
A natural starting point is to redo the NASD's 1943 analysis by asking what the size of mark-ups are in the sample at the seventy-one percentile and the size of mark-ups at the forty-seven percentile. It was, after all, the seventy-one percentile value that formed the original basis for the NASD's 5% guideline on mark-ups. The size of the mark-up at the seventy-one percentile in the sample is 2.95%, while the size of the mark-up at the forty-seven percentile is 1.8%. In other words, even for a sample heavily focused on less-liquid, lower-priced securities, the average mark-up is approximately 40% less than it was in 1943 at the same percentile cut-off points.

This is not at all surprising as the securities markets, from the less liquid to the heavily traded, are far more competitive and efficient, by any measure, than was the case some sixty-five years ago. Charles Jones, for instance, computed various measures of stock market liquidity and trading costs for NYSE securities over the course of the twentieth century.[29] The average annual turnover (defined as annual share value traded divided by total shares listed on the NYSE) rose from an average of 16% in the 1940s and 1950s to an average of 88% in 2000. The average one-way transaction costs (half the spread plus commission costs) of a trade fell from the 0.8% to 1.1% range to approximately 0.2 - 0.3% by the late 1990s.

[29] Charles Jones, A Century of Stock Market Liquidity and Trading Costs (2002).

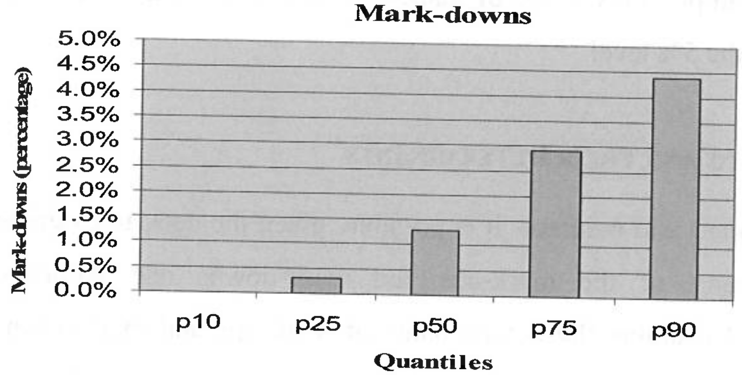
Mark-ups at the ten percentile (P10), twenty-five percentile (P25), median (P50), seventy-five percentile (P75) and the ninety percentile (P90) for the sample are summarized in Figure 1 below.

Figure 1.
Mark-ups



The same exercise was undertaken for the size of the mark-downs in the sample, which is summarized in Figure 2 below.

Figure 2.
Mark-downs



Interestingly, the mean mark-up (2.2%) and mark-down (1.9%) are higher than the median mark-up (2%) and mark-down (1.3%). The median mark-up of 2% compares favorably to the median mark-up of 4.1% documented by the SPECIAL STUDY in 1964, which, like this study, focused on the pricing practices of retail, over-the-counter, small- to medium-sized broker-dealers. The difference between the mean and median mark-up and mark-down in the sample suggests that there are some particularly large mark-ups and mark-downs that are driving the mean away from the median value. The average mark-up, weighted by the size of the trade, is somewhat lower than the mean mark-up of 2% at 1.6%.[30] The weighted average mark-down at 1.2% is likewise lower than the mean mark-down of 1.9%. This divergence suggests that larger trades occur at lower mark-ups and mark-downs. This is consistent with the SPECIAL STUDY's 1964 study of mark-ups which found that "higher mark-ups appear to have occurred in the smaller dollar transactions."[31] Also apparent from these numbers is the fact that mark-ups tend to be larger than mark-downs, again consistent with findings documented by the SPECIAL STUDY.[32]

The range of mark-ups in the sample is wider than that of mark-downs with mark-ups ranging from -1.5% to 225% compared to a range of -3.4% to 99% for mark-downs. The increased presence of extreme outliers in the mark-up distribution compared to that of the mark-down distribution can be seen in their respective kurtosis. The kurtosis of the mark-up distribution exhibits far more kurtosis (at 1,505) than the kurtosis of the mark-down distribution (at 340). This pattern, however, did not carry over to mark-ups and mark-downs greater than 5% (a cut-off that was chosen given the NASD's 5% guideline). Of all the mark-ups in the sample, only 1.2% of mark-ups are greater than 5%, while 1.7% of mark-downs exceed the 5% level.

VI. MARK-UPS, MARK-DOWNS AND THEIR DETERMINANTS

Besides looking at means and medians, it is possible, given the data, to examine more closely the determinants of the mark-ups and mark-downs one observes. Specifically, this section will examine the determinants of mark-ups and mark-downs over the entire range of observed mark-ups and mark-downs and, more narrowly, the determinants of mark-ups and mark-downs that exceed the 5% level. The entire distribution, including mark-ups and mark-downs less than 5%, is of regulatory interest

---

30 The Booz-Allen Study conducted in 1965 documented a weighted mark-up of 1.4% for over-the-counter stocks for a sample that included a large number of liquid securities. See SEC Report on Public Policy Implications of Investment Company Growth, Chapter 5, H. R. Rep. No. 2337, 89th Cong., 2d Sess. (1966)

31 SPECIAL STUDY at 625; Table VII-20.

32 Id. at 626 ("In sales on a principal basis the service charge or mark-down is usually lower than in the case of mark-ups on the purchase side.")

as it can help identify the factors that drive mark-ups and mark-downs. Examining the determinants of mark-ups and mark-downs that exceed 5% is of special interest as these constitute outliers in the sense of exceeding the NASD's guideline. The determinants of mark-ups and mark-downs will be examined separately as the impact of various factors can depend on whether one is dealing with a customer purchase or customer sale.

A. Determinants of Mark-Ups

Turning first to the determinants of mark-ups, the dependent variable will be the mark-up in percentage terms. The independent variables are the trade volume in the security on the day of the transaction; shares outstanding; the number of shares being bought or sold; the price of the security; the total value of the transaction (number of shares being bought or sold * price); stock price volatility; whether the transaction was a principal transaction or not; and the bid-ask spread of the security in percentage terms.

These variables have the benefit of including ones identified by the SPECIAL STUDY as factors affecting mark-ups and mark-downs.[33] On a similar note, the NASD likewise explained that "factors which are to be considered . . . in connection with the application of the [mark-up] policy" include "whether it is a low price stock or bond or high priced"; "the amount of funds involved in the transaction;" and "whether the market is active or inactive."[34]

Interestingly, four of these variables had no statistically significant positive or negative correlation with the size of the mark-up charged: the trade volume in the security on the day of the transaction; shares outstanding; the number of shares being bought or sold; and the total value of the transaction. The other four variables all had a statistically significant association (at the 1% level) with the size of the mark-up. The values of the coefficients on these variables, and their associated standard errors, are summarized below in Table I. While the four non-significant variables were included as independent variables in the regression, their coefficient values are not reported in Table I given their lack of significance.

---

[33] Id. at 624-625 (identifying principal or agency status, whether the purchase is a sale or purchase, the size of the transaction and stock price as explanatory factors).

[34] The identification of these factors can be traced to the NASD's May 16, 1949 statement. N.A.S.D. 11-15-55, E-135.

Table I. Regression result for Mark-ups (Percentage) as Dependent Variable

Independent Variable	Co-efficient	Standard Error
Price	-0.02	0.00
Volatility	0.01	0.01
Principal	1.31	0.01
Bid-ask Spread	-0.01	0.00

As is evident from Table I, by far the most economically significant coefficient value is whether the broker-dealer was acting as a principal or not in the trade. Principal transactions, not surprisingly, are associated with a substantially larger mark-up of some 131 basis points. This is not surprising as one would expect principal transaction to involve greater risk for the broker-dealer.

The signs of the coefficient values for price and stock price volatility also have the sign one would expect: an increase in stock price reduces the mark-up charged while an increase in stock price volatility increases the mark-up. One interpretation of these results is that stock price is a proxy for liquidity (lower-priced securities tend to be less liquid, especially in the context of a sample, as here, with a median price of $5.50), while stock price volatility acts as a proxy for risk. Not too much should be made, however, of the importance of the stock price volatility measure as it has marginal economic importance, at least within the sample. A move from the tenth percentile (P10) in terms of stock price volatility, i.e. a security with comparatively low levels of volatility, to a stock at the ninety percentile (P90) in terms of stock price volatility increases the mark-up charged by a mere basis point. The bid-ask spread is of even less importance affecting mark-ups by a mere 1/100 of one basis points when one moves from P10 to P90 in terms of the size of the bid-ask spread. In contrast, the stock price variable is of substantially greater economic importance than either the bid-ask spread or the stock price volatility variable: an increase in the stock price from P10 to P90 results in a 32 basis point reduction in the mark-up.

As a follow-up to the regression analysis, a probit regression was run with the dependent variable being whether the mark-up exceeded 5%. The same overall results appeared, with the coefficient values for trade volume, shares outstanding, the number of shares being bought or sold; and the total value of the transaction being insignificant. A higher stock price was associated with a lower probability of being charged more than a 5% mark-up. Interestingly, while a principal transaction is associated with a significantly higher mark-up, the probability of being charged more than 5% is slightly reduced (although the economic importance is marginal). In other words, agency transactions are associated with a somewhat increased likelihood of being charged a mark-up greater than 5% even though the average mark-up for agency transactions is lower than for principal transactions.

B. Determinants of Mark-downs

Roughly half the sample consisted of customer sales. Accordingly, a similar analysis was conducted on the determinants of mark-downs (as always in percentage terms) using the same independent variables. Once again, trade volume; shares outstanding; the number of shares being bought or sold; and the total value of the transaction had no statistically significant association (at the 10% level) with the mark-down charged customers. In addition, the bid-ask spread was statistically insignificant. The other three variables, as was the case with mark-ups, all had a statistically significant association, at the 1% level, with the size of the mark-down. The values of the coefficients on these three variables, and their associated standard errors, are summarized below in Table II. While the five non-significant variables were included as independent variables in the regression, their coefficient values are not reported in Table II given their lack of significance.

Table II. Regression result for Mark-downs (Percentage) as Dependent Variable

Independent Variable	**Co-efficient**	**Standard Error**
-		
Price	-0.02	0.00
Volatility	0.06	0.01
Principal	1.38	0.02

As with mark-ups, whether the transaction was a principal transaction or not had a substantial effect, with statistical significance, on the size of the mark-down quite similar to the magnitude of the effect of principal status on mark-ups charged. Stock price volatility had only a modest effect on the mark-down, increasing the mark-down by 6 basis points as one moved from P10 to P90 in the sample. Stock price also had a noticeable, although hardly overwhelming, effect on the mark-down charging being responsible for a 7 basis increase in the mark-down as one moved from P10 to P90. As with stock price volatility, the effect of stock price on mark-downs is quite similar to that documented for mark-ups.

When a probit regression was run to see which variables affected the probability of being charged more than 5%, the results track the findings just described. Stock price volatility has little effect while an increase in stock price reduces somewhat the probability of being charged more than 5%. Unlike mark-ups, principal transactions are associated both with higher mark-downs and a greater probability of being charged 5%.

C. *Stock Price*

Stock price, both in the ordinary regression analyses and in the probit regressions, is an important factor affecting the size of mark-ups and mark-downs. This is consistent with the reliance on stock price as a legitimate factor affecting broker-dealer prices in the decisional law on mark-ups. In particular, a common claim is that stocks selling for below $10 per share can warrant the broker-dealer charging customers' higher mark-ups and mark-downs than would otherwise be the case. Given the perceived importance of stock price as an explanatory factor, this variable merits further investigation.

The stock price appears to have little effect on the mark-up or mark-down charged once stock price exceeds, roughly speaking, $5.00, but is strongly correlated with the mark-up or mark-down charged for prices below this amount. This calls into some question whether $10 is the right price breakpoint for when the price level of the stock becomes an important explanatory variable. The SPECIAL STUDY attempted to answer the question of the proper price breakpoint by looking at the effect of stock price on the mark-up charged for 328 transactions in 53 over-the-counter stocks. However, the SPECIAL STUDY's analysis had very few transactions in which the price of the security was less than $10 making it quite difficult to draw any conclusions for transactions involving securities in this price range (with 19 transactions involving securities worth less than $5 and 18 transactions involving securities worth between $5 and $10 dollars).[35]

The non-linear relationship between stock price and mark-up charged raises the following question: which price breakpoints best explain the size of the mark-up/mark-down charged? After experimenting with a large number of break-points, the following breakpoints appear to be inflection points for the size of both mark-ups and mark-downs: 40 cents, $1, $5 and $10. The effect on mark-ups, for example, of having a price greater than 40 cents as compared to being less than 40 cents is, on average, 320 basis points. The effect of these price breakpoints was estimated by creating dummy variables for each price breakpoint and measuring their coefficient values with the dependent variable being the size of the mark-up or mark-down (as always in percentage terms). The association between these price breakpoints and the size of the mark-up or mark-down charged is summarized in Table III below.

TABLE III: PRICE BREAKPOINTS

Price Breakpoint	Mark-up	Mark-down
40 Cents	-320 basis points	-385 basis points
$1	-96 basis points	-78 basis points
$5	-28 basis points	-20 basis points
$10	+ 9 basis points	-7 basis points

Table III indicates that much of the association between stock price and mark-ups/mark-downs is in the lower price ranges with the price breakpoint of 40 cents representing by far the most important threshold. The effect on the mark-up or mark-down charged of whether or not the security being bought or sold is above or below the five dollar mark and, even more so, above or below the ten dollar mark is relatively unimportant (the size of the effect at the $10 level is quite small). This suggests that the price breakpoints should be lowered considerably from the $10 that is sometimes mentioned in the mark-up context with much of the focus being rather on whether the stock in question is a penny-stock. The fact that what was once considered "low-priced

---

[35] SPECIAL STUDY, Table VII-21.

securities" for purposes of determining whether a mark-up was excessive should be considerably tightened is consistent with the observation that securities markets today are far more efficient and competitive than they were several decades ago.

D. Solicitation

One of the interesting pieces of data gathered by the NASD from the fourteen broker-dealers was whether the transaction was solicited by the broker-dealer. Indeed, approximately 80% of all the transactions in the sample were in fact solicited.

The solicitation status of a transaction is of interest in its own right for at least three reasons. First, the law has long viewed broker-dealers as often operating under a heightened obligation to the customer when the transaction is solicited. For instance, the *Hughes* court, in the course of concluding that there was an implied broker-dealer representation that the price charged bears a reasonable relationship to the prevailing market price, pointed to the solicited nature of the transaction at issue. Second, there has long been the concern that solicited transactions increase concerns over investor protection. Is the transaction being solicited because it represents an attractive deal to the customer or for some other, perhaps more self-serving, reason? Third, there is reason to believe that solicitation status might be a good candidate as an explanatory variable for the price charged given that one would expect the costs of solicitation (such as hiring a sales force) to be recaptured through increased mark-ups and mark-downs. Indeed, one explanation for why mark-ups tend to be larger than mark-downs relates to the fact that broker-dealer sales tend to be solicited transactions a higher percentage of the time than broker-dealer purchases.

A transaction that is solicited, controlling for all the other independent variables examined in Sections A and B, is associated with a 57 basis point increase in the mark-up charged, but only an 8 basis increase in the mark-down. This is therefore a substantial difference between mark-ups and mark-downs in terms of the impact of solicitation. As a result, there might well be a basis for greater regulatory concern regarding solicited broker-dealer sales than there is for solicited broker-dealer purchases.

E. "Problematic" Trades

Only a small percentage of mark-ups and mark-downs in the sample exceeded the NASD's 5% guideline. Approximately 1.7% of all mark-downs exceed 5% while only 1.2% of mark-ups exceed this threshold. In comparison, 10.6% of transactions that the SPECIAL STUDY examined had mark-ups of more than 5%.[36] The dramatic reduction in the incidence of 5%+ transactions in the last forty-years is consistent with the development of more efficient and competitive capital markets, even for less liquid and lower priced securities.

As described in the data description, the NASD received, among other items, information from the fourteen broker-dealers concerning their standard commission schedules. Approximately, 21% of broker-dealer sales to customers that were in excess of 5% were also in excess of the charge indicated on the standard commission schedule for the type of transaction in question. Some 26% of broker-dealer purchases from customers that were in excess of 5% were also in excess of the charge indicated on the standard commission schedule. Trades in which the mark-up mark-down both exceeded 5% and the charge indicated on the standard commission schedule were labeled "problematic." These trades were labeled as problematic as they exceeded the NASD 5% guideline as well as the proper charge as indicated by the broker-dealer's own internal pricing schedule.

The question that was then explored was which factors are correlated with "problematic" trades. Probit regressions were run in which the dependent variable was whether the trade was "problematic" so defined. Several findings from this analysis are worth highlighting. As one moved to more and more expensive trades, as measured by the size of the charge indicated on the broker-dealer's standard commission schedule, both mark-ups and mark-downs were more likely to be "problematic." Interestingly, the effect of an increase in the charge indicated on the standard commission schedule on the likelihood of a trade being "problematic" was significantly greater for mark-downs than it was for mark-ups. As was documented earlier, principal status is an important determinant of the size of mark-ups and mark-downs. Interestingly, the broker-dealer acting as a principal made it more likely that there is a "problematic" mark-up as

[36] Id. at Table VII-20.

compared to the increased likelihood of a "problematic" mark-down. Finally, consistent with the analysis of the determinants of mark-ups and mark-downs, an increase in stock price rendered it slightly less likely that a trade is "problematic."

VII. CONCLUSION

The law of mark-ups has long focused on which factors properly drive the size of mark-ups and mark-downs in the course of determining whether a particular mark-up or mark-down should be deemed "excessive." It is difficult to answer this question on a case by case basis without the benefit of a large sample of transactions. As a result, this paper has investigated, using a unique dataset, the determinants of mark-downs and mark-ups. In particular, solicitation status appears to be especially important for the size of the mark-up charged, principal status for both mark-ups and mark-downs, and various price breakpoints, well below the $10 level, being quite important for both mark-ups and mark-downs. The paper also documents the substantial reduction in broker-dealer mark-ups and mark-downs over the course of the last forty years (since the SPECIAL STUDY examined the determinants of mark-ups and mark-downs) even in a sample largely consisting of lower-priced, less liquid securities.

39

Old Age and the Decline in Financial Literacy

Abstract

We investigate whether knowledge of basic concepts essential to effective financial choice declines after age 60. Consistent with prior studies of cognitive decline in old age, we find that financial literacy scores decline by about one percentage point each year after age 60. We test for possible cohort effects and find that the rate of decline in financial literacy is nearly identical among men, stockowners, older, and college-educated respondents. Confidence in financial decision making abilities does not decline with age. Older respondents with lower financial literacy scores are more likely to pay a high mortgage interest rate and less likely to capture credit card rebates. A separate analysis using data that include measures of cognitive ability and financial literacy suggest that a natural decline in both fluid and crystallized intelligence in old age contributes to falling financial literacy scores.

Michael Finke
Texas Tech University
Michael.Finke@ttu.edu
806-543-6724

John S. Howe
University of Missouri

Sandra Huston
Texas Tech University

1. Introduction

Households age 60 and over hold 51% of all financial wealth in the United States[1]. The proportion of U.S. households over the age of 60 is predicted to increase as the baby boom cohort and greater longevity contribute to population aging (Day, 2010). The transition to defined contribution plans tasks older Americans with greater responsibility for managing their own retirement assets and employing distribution strategies (Butrica, Iams, Smith, and Toder, 2009). Despite the importance of sound financial decision making among older Americans, little is known about either the magnitude of financial literacy decline in old age or about possible cognitive drivers of decision making quality.

There is evidence that financial decision-making ability declines in old age. Agarwal, Driscoll, Gabaix and Laibson (2009) show that the quality of credit decisions among borrowers erodes after peaking in the mid-50s. Despite evidence of improved general investment skill with age, investment performance declines significantly after age 70 (Korniotis and Kumar, 2011). Decision-making skills closely related to financial literacy such as the reliance on decision rules and resistance to framing decline in old age (de Bruin, Parker and Fischhoff, 2012). Hibbard, Slovic, Peters, Finucane and Tusler (2001) find that Medicare beneficiaries are nearly three times more likely than younger subjects to make errors when interpreting health plan information despite having more experience.

The observed decline in financial decision-making quality may be related to gradual mild cognitive impairment that occurs in old age. Boyle et al. (2012) find that the rate of cognitive decline in a sample of older adults is a significant predictor of incorrect responses to a financial decision-making test and increases susceptibility to financial scams. Older financial decision makers who experience a sharper decline in cognition report an increased difficulty in managing their money (Hsu and Willis, 2013). Cognitive ability, and in particular mathematical skills of the primary financial decision maker, is a strong predictor of the ability to avoid depleting net worth in later life (Smith, McArdle and Willis, 2011) and in making fewer financial mistakes (Agarwal and Mazumder, 2013).

Horn and Cattell (1967) attribute the declining performance on mathematical or spatial reasoning tasks after young adulthood and improved performance on tasks that require experience and knowledge to the theory of fluid and crystallized intelligence. Large cross-sectional analyses of performance on decision making tests such as word recall are indeed highest for respondents in their 20s and decline gradually through middle age before falling sharply after age 60, but scores on vocabulary tasks such as the ability to produce a synonym peak late in life and begin to fall after age 60 (Salthouse, 2009). Fluid intelligence decline appears to be the result of a general slowing in cognitive processing ability (Bugg,

[1] Calculated from the 2008 Survey of Consumer Finances.

Zook, DeLosh, Davalos and Davis, 2006) that has been linked to physiological changes such as a decline in frontal lobe volume after age 50 (DeCarli et al., 2005, Rushton and Ankney, 2009).

Financial literacy can be defined as the ability to understand fundamental financial concepts needed to make effective decisions. It requires an understanding of terminology, for example, a deductible on an insurance policy or the characteristics of a mutual fund, and the ability to comprehend how a higher deductible lowers an insurance premium or how greater diversification is a benefit of a mutual fund. Studies on information retrieval indicate that the ability to recognize terms may not decline in old age, but there is evidence that interpretation and general problem solving capability deteriorates. For example, Salthouse (2010) finds no significant decline in subjects' ability to solve crossword puzzles after age 60. Performance on more complex tasks that require the ability to retrieve information and use that information to solve a problem appears to worsen in old age. The ability to complete everyday tasks, for example to read and interpret instructions on a medicine bottle or interpret a rate chart on a telephone bill, decreases after age 60 (Diehl, Willis and Schaie, 1995). de Bruin et al.(2012) find that decreasing scores on fluid cognitive ability tests in old age contributes to the decline in performance on decisions that require both problem solving and accumulated knowledge. The ability to retrieve financial terms may or may not decline with age, but the ability to make an appropriate financial choice is particularly vulnerable to the age-related decline in reasoning skills.

Previous studies suggest a possible decline in financial literacy in old age. Lusardi, Mitchell and Curto (2014) identify differences in financial sophistication among older respondents in a subsample of the 2008 Health and Retirement Study. While the authors focus on how demographics impact knowledge scores, they note that respondents over 75 are less likely to understand basic investment concepts such as stock diversification and the importance of mutual fund fees. Consistent with a loss in fluid intelligence, respondents over 75 also score lower on numeracy questions. Lusardi, et al.(2014) do not investigate the rate of decline during old age nor does it investigate whether the decline is related to demographic differences among older cohorts. Descriptive results from van Rooij, Lusardi, and Alessie (2011) show that the proportion of respondents in the highest basic financial literacy quartile peaks in the 41-50 age category and is lowest among those ages 71 and older. The inverted U-shaped relation between age category and literacy is unexpected because rates of stock ownership and net worth are both related to higher financial literacy scores and rise with age. In multivariate analyses, van Rooij et al. (2011) do not segment respondents over the age of 60 and do not find that age is a significant predictor of financial knowledge.

Higher financial literacy scores have been linked to higher quality financial decisions. Hilgert, Hogarth and Beverly (2003) find that higher financial literacy scores predict on-time credit repayment, investment diversification, and mortgage refinancing. Lower financial literacy is associated with incurring fees that are 50% higher on credit cards, particularly fees that require a more sophisticated awareness of credit terms (Lusardi and Tufano, 2009). A particularly costly financial mistake for older households in a falling interest rate environment is the failure to refinance a mortgage. Campbell (2006) finds that characteristics associated with financial sophistication predict wealth-maximizing refinancing behavior.

Our study adds to the existing literature on financial decision making and age by estimating the actual rate of decline in financial knowledge and measuring the decline within subgroups to address possible cohort biases. We also estimate the relation between the decline in fluid and crystallized intelligence and financial literacy scores, and investigate whether age-related financial literacy declines can affect financial decision-making quality. Studies that provide evidence of a decline in investment performance with advanced age (Korniotis and Kumar, 2011) and in credit decision making (Agarwal et al., 2009) do not directly estimate the decline in financial literacy that may be driving reduced performance in decision making ability. These performance studies also do not use data that allow them to accurately control for important demographic characteristics such as gender, race, and education that may influence observed behavior through differences in mortality rates or cohort differences in human capital investment. We extend the analysis of Lusardi et al.(2014) by using a much larger population of older households that allows subgroup analyses to address cohort and gender biases, and we model the rate of decline in financial literacy with advanced age.

We use a new financial literacy assessment instrument inserted into the Consumer Finance Monthly (CFM), a nationally representative monthly survey of credit behaviors, and obtain a sample of 3,873 respondents over age 60. The financial literacy score is composed of four questions, each within the topic areas of basic financial concepts, insurance, investments and credit knowledge. This unique instrument allows us to assess total financial literacy as well as more specific knowledge in financial topic areas. Financial literacy may be lower among older cohorts because of less investing experience in the pre-401(k) era, because women generally outlive men and may have delegated financial decisions to husbands, or because levels of higher education were lower for older cohorts. We estimate performance among domains and focus on insurance knowledge because rates of insurance ownership are higher among older cohorts. Our large sample size allows us to estimate the decline in financial literacy scores among college graduates, stockowners, men, and cohorts who reached age 60 before the rise in 401(k) popularity.

The survey also includes a self-assessed measure of confidence within each of the four financial literacy domains, and we estimate whether confidence declines with age and whether those with lower financial literacy have lower confidence. In order to identify possible cohort knowledge biases among topic areas, we estimate whether financial literacy declines with age for each of the 16 questions. Using available financial decision quality variables, we estimate whether financial literacy scores predict interest rates on mortgages, observed understanding of financial questions, and the use of reward credit cards. Using the Health and Retirement Study, we create a financial literacy instrument and model scores as a function of age and measures of both fluid and crystallized intelligence to test whether age-related changes in cognition are associated with a decline in financial literacy scores.

We find a consistent linear decline in financial literacy scores after age 60 and a monotonic decline in scores among 5-year cohorts in the CFM data. Performance on all 16 questions declines significantly with age. The annual rate of decline in financial literacy scores is significant and similar among all subgroup analyses of men, stockowners, those with a college degree, and respondents who were 60 or older by 1992. The magnitude of age-related decline in scores is nearly identical among all four financial topic areas, including insurance. Confidence in one's ability to make financial decisions does not decline in old age and increases significantly for insurance. Age is positively related to financial overconfidence measured as the distance between objective literacy score and subjective financial confidence. Financial literacy scores predict interest rates on mortgages, observed understanding of financial questions, and the use of reward credit cards. Using data from the Health and Retirement study, we find that lower financial literacy scores within older age cohorts are related to the decline in fluid and in crystallized intelligence.

2. Methods

2.1. Financial Literacy Assessment

Some studies have measured financial literacy subjectively (e.g., Stango and Zinman, 2014) by asking respondents to estimate their perceived level of financial knowledge and capability. Other studies use measures that primarily measure numerical skills (Lusardi and Mitchell, 2007a, 2007b, Gustman, Stenmeier, and Tabatabai, 2012). Numeracy, or math capability, does not necessarily equate with having made the investment in knowledge of financial concepts needed to make optimal financial decisions (Almenberg and Widmark, 2011; Banks, 2010; Hung, Parker and Yoong, 2009). Other financial literacy measures focus on knowledge by asking respondents to recognize the correct definition of a financial term (Moore, 2003; Hilgert et al., 2003) while others attempt to incorporate skill by including items that require respondents to appropriately apply their knowledge (Mandell, 2008; Lusardi and Tufano, 2009).

Even among the measures that include financial knowledge and skill, none comprehensively cover the four main topic areas of financial literacy (Huston, 2010). Financial knowledge and skills encompass a broad array of topics ranging from awareness of financial terms to an understanding of financial concepts and the ability to select appropriate financial instruments. The objective of the financial literacy instrument used in this study is to measure both financial knowledge and the ability to apply knowledge effectively. A research team initiated development of a financial literacy instrument, tested 89 potential questions and analyzed responses with the goal of choosing questions that were not biased in terms of age, gender, race and socioeconomic status, had an unambiguous correct response, and correlated well with other high quality questions. An eight-member panel of national experts in financial literacy and its assessment reviewed the project (including goals, design, model, instrument, scoring, and results) in order to assess the proposed methods and assessment instrument.

The final instrument of 20 items is selected from the best performing financial literacy questions according to reliability and validity statistics and the recommendations of the expert panel[2]. In December, 2009 through 2013, the 20-item survey was included as a module in the Consumer Finance Monthly Survey conducted by the Center for Human Resource Research at The Ohio State University. The Consumer Finance Monthly collects demographic and detailed credit use information through a random digit dialing phone survey in the United States. The survey began in 2005 and includes over 10,000 completed financial literacy assessment instruments.

The sample in this study includes financial literacy responses from 3,873 respondents age 60 and older during this time period. The financial literacy instrument contains 20 items (Appendix B) covering the four content areas of personal finance – basics, investments, credit and insurance. Analyses show the questions within the instrument show high internal consistency[3]. Within each of those four personal finance content areas, there are two knowledge questions, two ability questions, and one confidence question. There are 16 questions used to measure the objective financial literacy score and four questions that measure confidence. Basic personal finance concepts include elements such as time value of money, purchasing power, and personal finance accounting. Intertemporal transfers of resources include both borrowing (bringing future resources into the present for consumption through the use of revolving credit and installment loans) and investment (saving present resources for future consumption through the use of savings accounts and investing through stocks, bonds, or mutual funds). Insurance questions include insurance instruments and risk management techniques. Financial literacy score is estimated as the percent correct out of 16 questions or out of four questions when scores are calculated within each topic area (basics, borrowing, investing and insurance). Decision making confidence is measured on a scale of 1 to 10 in each topic area and in total from 4 to 40 for the four topic areas.

The individual financial literacy questions require an understanding of basic financial products and an ability to apply them appropriately. In this sense, they test both knowledge of financial products

---

[2] The complete financial literacy assessment instrument can be found at https://sites.google.com/site/pfinttu/flat

[3] Construct validity estimates for our financial literacy assessment instrument are higher than for previous financial literacy instruments, and the sample size is more than twice as large as any previous literacy module (Hung et al., 2009). Cronbach Alpha estimates and correlations among questions can be found in Appendix A.

that will likely improve from age and experience, and some reasoning skill which may decline in advanced age. For example, an insurance question asks what impact a higher deductible will have on an insurance premium. This requires an understanding of the financial terms deductible and premium, and the ability to think through how a higher insurance deductible will affect the cost of insurance.

2.2. Measuring the Decline in Financial Literacy in Old Age

Our first objective is to test whether financial literacy scores decline among respondents age 60 and older, and to estimate how the rate of line changes in advanced age. We then test whether respondent characteristics other than age are associated with financial literacy. The greatest challenge to any cross sectional analysis of knowledge assessment is the possibility of cohort effects that may create estimation biases. We conduct a number of subgroup analyses in order to test whether the hypothesized negative relation between age and financial literacy remains consistent.

To create Table 1, we calculate average financial literacy score for each year of age among the 3,873 respondents age 60 through 94 in the CFM for a total of 35 years. For example, there are 265 respondents age 60 and the average financial literacy score is 61% among 60-year old respondents. Cross sectional estimation of the marginal change in average test score for each year of age is common in the cognitive aging literature (Salthouse, 2010)[4]. We model average score as the dependent variable and age as a single independent variable. Regressions are estimated using the average overall 16-question financial literacy score as a dependent variable, the overall 4-question confidence score, scores within each of the 4-question topic areas (basics, borrowing, investments and insurance), and confidence within each of the 4-question topic areas. Average financial literacy score for all ages is included in Figure 1, and Figure 2 shows the average decline in financial literacy and average confidence score by age. Figure 3 shows average score within each topic by age.

2.2.1. Controlling for Respondent Characteristics

In the multiple regression analyses, we model financial literacy as a function of demand for financial human capital. The decision to incur the direct and indirect costs of attaining financial

---

4 Although there is some debate about the use of cross-sectional data to identify age-related decline in task ability, Salthouse (2009) illustrates how longitudinal estimate biases caused by subject learning create significant problems in panel data. Even questions that measure special orientation and word recall show increased ability in subsequent panel surveys among young and old respondents, while cross-sectional data show a consistent decline (Salthouse, 2009). The best method of estimating age-related task decline is to carefully reduce potential cohort effects through empirical models that control for differences among age groups that may be related to task ability.

knowledge is a function of the time and transaction costs and the discounted expected utility from making more effective financial decisions in future periods.

Higher education may proxy a lower cost of information acquisition, a lower rate of time preference, or may involve direct exposure to financial information via business or economics coursework – all of which will increase expected financial literacy. Home ownership may be related to financial literacy both through experience with related financial products (for example insurance concepts), and by increasing the expected return to investment in tax rules. Likewise, stock ownership may involve a fixed information cost that suggests a greater expected benefit from investment in financial human capital (Peress, 2004). The use of tax-sheltered accounts requires an initial financial human capital investment and may help explain greater investment knowledge among those who actively saved during the 1980s and 1990s when the use of sheltering instruments expanded in the U.S. We use a question that asks respondents whether they have "any money in tax advantaged accounts including IRAs, Keogh plans, variable annuities or 529 plans" or "money in retirement plans through former employers such as a 401(k) or 403(b)."

Financial wealth will increase the expected future payout from investing time and effort into making more informed financial decisions (Peress, 2004). We use the top income and wealth quintile to capture the incentive to invest in financial information among those with the most money to manage. To some extent, home ownership, stock ownership, the ownership of tax sheltered accounts and marital status will also capture financial resource availability. Racial differences in financial literacy may be attributable to differences in financial human capital inherited from parents or to differences in the frequency of financial knowledge transfer in social interactions (Brown, Izkovic, Smith and Weisbenner, 2008). Women may have lower financial literacy if households allocate financial decisions to the spouse with a lower relative cost of financial capital acquisition (Croson and Gneezy, 2009). Smith et al.(2010) find that women are less likely to be the primary household financial decision maker in older household cohorts. We also include dummy variables for the year and month of the survey.

In order to estimate the impact of age on financial literacy among individuals, we model the percent correct on financial literacy topic areas and total score (percent correct of 16 questions) as a function of age and control variables using an ordinary least squares (OLS) regression:

$$FinLit_i = a + bAGE_i + \gamma X + \varepsilon_i \tag{1}$$

where $FinLit_i$ is the objective financial literacy score (percent correct out of 16 questions) for respondent i and AGE is the respondent's age in years. We include a matrix of control variables (X) that yield a vector of coefficient estimates (γ). The control variables include education, income, and wealth levels, along with gender, race, marital status, and ownership status of home, tax shelters, and stocks (see equation 1).

In a second regression (see equation 2), interaction variables of age measured as a continuous variable (*AGE*) multiplied by whether the respondent is age 60-69, 70-79, or 80 years of age or older (*AGECAT*) are substituted for a linear age variable to account for possible slope differences in the effect of age on financial literacy.

$$FinLit_i = a + bAGE_i *AGECAT + \gamma X + \varepsilon_i \tag{2}$$

To provide a comparison of the independent age effect across the entire sample, we model total financial literacy score as a function of age (*AGE*) and age-squared (*AGE*^{2}) to capture the hypothesized inflection point of financial literacy in middle age (see equation 3).

$$FinLit_i = a + bAGE_i + cAGE^2{}_i + \gamma X + \varepsilon_i \tag{3}$$

We also specify age using 5-year age cohorts using age 45-49 as the reference category, and provide actual average financial literacy scores by age, to create Figure 1. For example, the age cohort scores are estimated based on the coefficients with control variables set to their mean value from the model:

$$\begin{aligned} FinLit_i = {} & a + bAGE25\text{-}29_i + cAGE30\text{-}34_i + dAGE35\text{-}39_i + eAGE40\text{-}44_i + fAGE50\text{-}54_i + \\ & gAGE55\text{-}59_i + hAGE60\text{-}64_i + jAGE65\text{-}69_i + kAGE70\text{-}74_i + lAGE75\text{-}79_i + mAGE80\text{-}84_i \\ & + nAGE85\text{-}89_i + oAGE90\text{-}94_i + \gamma X + \varepsilon_i \end{aligned} \tag{4}$$

2.2.2. Addressing Sample Biases

Differences in experiences or incentives to invest in financial knowledge may affect performance on the financial literacy test. To minimize possible biases in older cohorts, we create subsamples that reduce the most significant sources of age-related financial knowledge variation that are unrelated to cognitive decline.

Because rates of educational attainment have risen in the United States during the 20th century (Day, 2010), younger cohorts may be more likely to have taken an economics or finance course in college. To reduce the potential bias caused by lower educational attainment by older cohorts, we estimate our model only on older respondents who have a college or graduate school education. Men, particularly in older cohorts, may choose to invest in financial knowledge as a result of specialized labor in household production. Because longevity is higher among women than men, our results may be biased by a larger proportion of older, less knowledgeable women. To correct for gender-related knowledge differences, we estimate the model among males only. Cohort differences in rates of return on stock investments may drive variation in equity market participation (Malmendier and Nagel, 2011). If households over 60 are less inclined to invest in equities due to their poor performance in the 1970s, this may have influenced the decision to seek out investment information about stocks. To address the bias or reduced preference for risky asset ownership, we estimate the model using households who directly hold stock or mutual fund investments. Finally, since financial literacy is strongly related to stock market participation (van Rooij, Lusardi and Alessie, 2011), and household stock ownership rates peak among households age 45-54 and rose significantly between 1983 and 1992 in qualified retirement plans (Poterba and Samwick, 1995), we estimate our model only for households who were age 60 or older in 1992.

Another potential criticism of estimating the relation between age and financial literacy is the possibility that older households were less likely to be exposed to financial instruments less common in their peak borrowing and saving life cycle years. One exception is insurance products. Cohort ownership rates of life insurance are higher among older households than among the baby boomer cohort (Chen, Wong and Lee, 2003). In addition, insurance products were a common sheltered savings vehicle prior to the 401(k) and IRA era that began in the 1980s. The four financial literacy questions related to household insurance present less potential cohort bias than other topic areas. We estimate OLS regressions on the percent correct (out of four questions) using equation 1 from within each of the four financial literacy topic areas, including insurance, investments, borrowing and basics in order to detect possible differences in the marginal effect of aging against knowledge in different literacy domains.

We estimate sixteen separate logistic regressions using equation 1 on each financial literacy question to determine whether the results are driven by a subset of questions that may be age or cohort biased. We separately estimate standardized beta coefficients in order to calculate the relative independent strength of the age variable among independent predictors of financial literacy within the multivariate model.

2.3. Financial Confidence

The financial literacy assessment instrument includes four questions that ask the respondent to assess how confident they are at making financial decisions within each of the topic areas. It is possible that older subjects are not aware of declines in their financial decision making ability and may or may not remain confident of their financial capabilities. Understanding whether actual ability and age are related to higher confidence is important in understanding whether seniors are potentially vulnerable to decision making mistakes from overestimating their decision making ability. In order to better understand an over-confidence in one's abilities, we investigate the characteristics that predict a high level of confidence among respondents with a low level of knowledge.

Confidence in financial ability is measured through a question asking respondents to rate on a scale of 1 to 10 how confident they are in making decisions within each of the four financial literacy topic areas. We use an OLS model to estimate predictors of confidence in each topic area, and the summed total of all four topic areas, as a function of age, financial literacy within that topic area (or total financial literacy score), and household characteristic control variables:

$$Confidence_i = a + bFinScore_i + cAGE_i + \gamma X + \varepsilon_i \tag{5}$$

where $Confidence_i$ is either confidence in a specific topic area (basics, borrowing, investment, insurance) or total confidence (percentage total of all four topic areas) for respondent *i* (see equation 5). Confidence in each topic area is respondent-assessed on a scale of 0 (no confidence) to 10 (highest confidence). $FinScore_i$ is the specific topic area objective score (percent correct out of 4 questions) or the total score (*FinLit*) for respondent *i*. Coefficient estimates indicate the marginal effect of age on financial confidence controlling for actual ability. The sample size decreases slightly to 3,403 because of some missing responses among those who completed the financial literacy test.

Respondents whose confidence score is an average of at least 8 for all four topic areas (or a total score of at least 80%), and whose objective financial literacy score is in the lowest quartile, are considered overconfident. Since slightly higher confidence may improve financial outcomes for those with adequate financial literacy, we choose to specify overconfidence as very low financial literacy and very high confidence. We select a logistic model (see equation 6) of overconfidence as a function of age (*AGE*), and interaction of age and financial literacy (*AGE*FinLit*), and the same matrix of control variables (*X*).

$$Overconfidence_i = a + bAGE_i + cAGE{*}FinLit_i + \gamma X + \varepsilon_i \tag{6}$$

2.4. Financial Literacy and Decision-Making

Prior research shows a decline in financial decision-making quality in old age. We examine whether lower financial literacy is associated with an increased likelihood of selecting more efficient financial products and of understanding the questions posed in the survey. We also investigate whether overconfidence, independent of a lower overall financial literacy score, increases the likelihood of poor financial decision-making. In order to understand whether the negative impact of lower financial literacy scores and higher overconfidence differs among age cohorts, we include interaction effects of each variable with older age categories.

To gain an understanding of decision-making quality of older households, we create logistic regression models that include age categories, financial literacy, financial confidence and the age-score interactions, while including the matrix of control variables (see equation 7). We select two financial outcomes available in the CFM as the most unambiguous measures of decision-making quality among older households. The first outcome measures whether respondents with a mortgage have an interest rate higher than what is normatively optimal. Agarwal, Driscoll, and Laibson (2013) find that optimal refinancing differentials between existing mortgage and current interest rates are never greater than 200 basis points. We set a threshold of 6.5%, which is the lowest rate among respondents with a mortgage who fall in the highest rate quintile and roughly 200 basis points above the average rate during the survey period (4.46%). Campbell (2006) finds that variables related to financial literacy such as formal education are a significant predictor of owning a mortgage 2% above market rates. The second outcome measures whether the respondent is a convenience credit card user and does not take advantage of costless interchange fee reimbursement through a reward card (Ching and Hayashi, 2010). We censor to only respondents who own a credit card and who indicate that they pay off their credit card balance in its entirety each month, and thus do not consider card interest rate a salient characteristic. Retailers pay $30 billion in credit card interchange fees at a cost of $350 per household, and 44% of this (or $150 per household) amount is rebated to consumers through reward cards (Ching and Hayashi, 2010). The failure to capture this rebate among convenience users represents a significant loss arising from low financial literacy. The third outcome we model is whether the interviewer believes the respondent had a good understanding of the survey questions to test whether financial literacy predicts a global subjective assessment of financial decision making ability.

$$Outcome_i = a + bAGE70\text{-}79_i + cAGE80Plus_i + dFinLit_i + eOC_i + fFinLit^*AGE70\text{-}79_i + gFinLit^*AGE80Plus_i + hOC^*AGE70\text{-}79_i + jOC^*AGE80Plus_i + \gamma X + \varepsilon_i \tag{7}$$

where $Outcome_i$ is either the mortgage rate, credit reward, or interviewer's assessment status for respondent *i*. $FinLit_i$ represents the financial literacy score and OC_i represents the overconfidence status (1=overconfident; otherwise 0) for respondent *i*.

2.5. Financial Literacy Decline and Cognitive Aging

If the decline in financial literacy in old age is not explained by cohort differences or household characteristics, it is possible that cognitive decline in advanced age is responsible for reduced financial decision-making ability. The 2010 Health and Retirement Study (HRS) contains a module that asks financial literacy questions to a subsample of respondents. The HRS also asks respondents to answer questions in a cognition module that include established measures of fluid and crystallized intelligence. Through the use of cognition measures, we are able to test whether lower financial literacy scores are related to by cognitive aging.

Financial literacy questions are drawn from an HRS module of questions that assess respondent "financial sophistication and investment decision making." The HRS module includes questions that measure financial literacy and other concepts related to financial awareness. We select the questions from the module that measure financial literacy. These include the original three questions from Lusardi and Mitchell (2007a) and four additional questions that have a specific answer ("buying a single company stock usually provides a safer return than a stock mutual fund"), and we avoid questions with no specific correct response ("are you considering investing in the stock market for the next year?"). The financial literacy instrument is an additive score that includes one question on interest compounding and one on inflation (each of these may be related to numeracy), whether a single stock is safer than a mutual fund, whether stocks historically provide higher returns than bonds or savings accounts, whether an employee should have a lot of their retirement savings in employer stock, whether foreign stocks should be avoided, and whether bond values are inversely related to interest rates.

Using the percent correct from an additive score of 7 financial literacy-related questions as the dependent variable, we model using OLS financial literacy as a function of age and cognitive function while controlling for a matrix of household characteristics (*X*) for respondents age 60 and older (see equation 8). We estimate models that specify age as a linear variable and as 5-year cohorts using age 60-64 as the reference category (see equation 8).

$$HRS\text{-}Score_i = a + bAGE_i + cRecall_i + dVocab_i + \gamma X + \varepsilon_i \tag{8}$$

where $HRS\text{-}Score_i$ is the percent correct out of 7 financial literacy-related questions from the HRS for respondent i and AGE is the respondent's age in years (or age category using 5-year intervals). We calculate fluid intelligence using a combination of immediate and delayed word recall scores, a reliable measure available in the Health and Retirement Study (HRS) (McArdle, Fisher and Kadlec, 2007). $Recall_i$ is the word recall is the number of nouns (out of 10) recalled by respondent i. $Vocab_i$ is a vocabulary measure adapted from the WAIS-R crystallized intelligence test that asks respondents to define each of five words (such as plagiarize or perimeter). Responses are scored between 0 (wrong) and 2 (perfectly correct) for a total score of between 0 and 10. Analyses are performed using a sample of 1,109 respondents who answered the financial literacy module for the analyses without the cognition variables and a total of 887 respondents who completed both the cognition and financial literacy questions.

3. Results

3.1. Financial Literacy Decline in Old Age

Figure 1 shows the average financial literacy scores by age and the score predicted by the multivariate models that specify age as a quadratic variable and as 5-year cohorts using the full CFM sample. Actual and predicted scores using the two function forms show a clear concave relation between actual financial literacy score and predicted score controlling for other respondent household characteristics. Average financial literacy scores within each year of age increase up to roughly age 50, while the quadratic specification controlling for demographic characteristics peaks at age 49, and 5-year cohort model is highest in the age 40-44 group (although it is not statistically different from the 45-49 year old reference group). Within the cohort model, the first cohort to have scores statistically lower than the 45-49 year old reference group is age 65-69 (3% lower) and predicted financial literacy scores decline at a rate of between 5 and 7 percentage points for each 5-year later life cohort. The predicted financial literacy percentage score (when compared to respondents age 45-49) is roughly the same among respondents between age 25 and 29 (7.9% lower) as it is for respondents age 70-74 (8.3% lower) and falls to 35% lower for respondents age 90 or older.

Univariate regression analyses in Table 1 model the yearly change in average financial literacy scores among respondents between the age of 60 and 94. Regression results show that age is a strong and consistent predictor of financial literacy score. With each year of age after 60, the average score falls by 1.5 percentage points, and the relation is consistent (R-square of 0.96). The relation between age and average financial literacy score is consistent among the four decision making topic areas ranging from a 1.42% decrease each year within basic questions to 1.65% for investment questions. Although financial literacy scores decline with age, confidence in financial decision making does not. Confidence in financial decision-making ability increases slightly with age, but the relation is statistically significant only within the insurance domain.

Figure 2 shows the similar rate of decline in average financial literacy score in the CFM and in average word recall ability by age in the HRS. Although episodic memory and financial literacy score decline at roughly the same rate after age 60, confidence in financial decision making ability remains relatively unchanged with age. The percentage of overconfident respondents with high self-assessed ability and low objective literacy scores increase from about 10% in the 60s to higher than 30% among respondents over 85. Figure 3 illustrates the consistent decline in financial literacy score with age among all four topic areas.

Sample characteristics in Table 2 indicate consistently lower financial literacy scores in all topic areas among respondents age 70-79 (49%) and 80+ (32%) than among respondents age 60-69 (62%). Older respondents have lower average financial literacy scores in all topic areas including insurance. Scores are much higher among respondents with a college (61%) and graduate (66%) degree than among respondents with a high school (39%) or below high school (25%) education. Scores are higher among whites, men, homeowners, those who are married, stock owners, and increase monotonically with wealth and income quintile. Greater financial resource availability is associated with higher financial literacy scores. Older households are slightly more confident in their financial decision making abilities (Table

2b). Since average financial literacy score declines with age, it is not surprising that a higher percentage of respondents age 80 or above (19.3%) are overconfident than respondents age 70-80 (10.4%) and age 60-69 (4.7%). Financial confidence is only slightly higher among more educated respondents, but a much higher percentage of respondents with a high school or below high school education are overconfident.

Table 3 presents regression results that estimate the financial literacy score by individuals with and without control variables[5] among respondents 60 years and older in the CFM. Unlike Table 1, the regression model predicts individual financial literacy score for each of the 3,873 respondents as a function of the respondent's age (rather than average financial literacy score for all respondents that have the same age in Table 1). Because financial literacy will vary among respondents of the same age with different demographics and life experience, the fit of the relation between age and financial literacy is lower (R-square 0.17) but still statistically significant. Each year of age is associated with a 1.36 percentage point decline in the total financial literacy score. When age is sorted into 5-year groups, respondents age 70-74 have significantly lower financial literacy scores than respondents age 60-64. The predicted financial literacy score falls by between 6 and 9 percentage points with each older cohort group. Respondents age 90 or older score 41.7 percentage points lower on average than respondents age 60-64. The age effect declines slightly to 1.02 percentage points per year when control variables are included in the model, but the effect is no less consistent. The model with control variables explains 37% of the variation in observed financial literacy. Coefficients in the multivariate model also decline slightly in magnitude but remain significant when age is measured in 5-year groups and the fit is identical to the linear age model.

3.1.1. Addressing Sample Biases

Regressions in Table 4 model financial literacy score using the linear age variable within subsamples in order to correct for possible biases in experience or motivation to acquire financial knowledge. Among households with a college education, the magnitude of the decline in financial literacy score is slightly higher than in the full sample (1.10 percentage points per year). The annual decline in financial literacy is similar in a sample of men (0.96 percentage points per year) compared to a full sample of male and female respondents. Among stock or mutual fund owners, the annual decline in financial literacy is 0.98 percentage points and also statistically significant. Among the cohort of respondents who were age 60 or older in 1992, the magnitude of decline with age is 1.37 percentage points per year. The estimated linear decline in financial literacy is similar among subgroups.

[5] A quadratic age specification yields an insignificant age-squared coefficient that is near zero indicating no inflection point of age on financial literacy in the older sample.

Financial literacy regression results for individual topic areas in Table 5 show the marginal impact of age on predicted score within the four areas of financial literacy knowledge. The annual decline in financial literacy scores is consistent among all four topic areas, and the magnitude of the effect is comparable (ranging from a 0.94 percentage point decline per year for basics to a 1.10 percentage point decline in borrowing knowledge). All results are statistically significant. Insurance knowledge, which would increase with age if results are driven by cohort financial instrument familiarity effects, declines with age at roughly the same rate (0.96 per year) as basic financial knowledge. The consistency of the age decline among financial literacy topic areas can also be seen when we model the correct response to individual questions. Table 6 shows that the likelihood of providing a correct response to each financial literacy question declines significantly with age. Of the 12 control variables (including important human capital-related characteristics such as education, income race, gender), age is the strongest independent predictor of providing the correct answer for 11 of the 16 financial literacy questions. Of the remaining questions, age is the second and third strongest predictor. Interestingly, the weakest age effect is observed for a question that asks about appropriate mortgage types for a first-time home buyer. The strongest age effect occurs for the questions on the deductibility of interest and the use of money market accounts.

3.2. Financial Confidence and Decision-Making

Table 7 shows multivariate analyses of confidence in managing money, managing credit and debt, using investment products and using insurance, as well as overconfidence measured as the difference between self-assessed and actual financial literacy. Confidence in overall financial decision making ability increases with age, and also within all topic areas. More financially literate respondents are also more confident for each topic area except insurance. Respondents who are less knowledgeable about insurance are not less confident about their insurance knowledge. Older respondents are more likely to be confident about their ability to make insurance and investment decisions.

The likelihood of being overconfident with one's financial knowledge increases with age. Each year of age after 60 increases the likelihood of having high confidence and low financial literacy scores by 7 percent. Higher levels of education are associated with a much lower likelihood of overconfidence, as are being male and white.

The only variable that consistently predicts confidence in all four areas is homeownership. Although age is related to increased financial confidence, the multivariate models explain little variation in financial confidence and the marginal effect of age is weak compared to other variables (the effect of an additional 20 years of age is roughly equal to homeownership in predicting total confidence).

Estimation of the impact of age and financial literacy on decision quality outcomes available in the CFM are presented in Table 8. Respondents with a higher financial literacy score are more likely to have a competitive mortgage rate beneath a threshold at which a borrower should refinance. There is no independent relation between overconfidence and a low mortgage rate. Like Campbell (2006), we find that education and wealth are significant predictors of a low mortgage rate. More financially literate and older convenience credit card users were also more likely to have a reward card. Again, there is no relation between overconfidence and reward card ownership. More financially literate respondents were more likely to have a good understanding of the survey questions according to the interviewer. Age and overconfidence had no independent impact on understanding.

3.3. Financial Literacy Decline and Cognitive Aging

Table 9 shows results using data from the Health and Retirement Study that include measures of respondent fluid and crystallized intelligence as well as a different measure of financial literacy. Coefficients represent the percentage change in financial literacy score out of seven questions included in the HRS.

The HRS financial literacy questions show a consistent, but slightly weaker, decline in financial literacy among respondents 60 and older. The linear age specification is negative and statistically significant, and the 5-year age group coefficients are negative and monotonic but only reach the level of a statistically significant difference from age 60-64 by the 75-79 age group. Similar to the model using CFM data, controlling for household characteristics slightly reduces the linear age effect. The magnitude of financial literacy decline is similar among 5-year age groups who have reached at least age 80 after demographics are controlled.

Both the measures of crystallized and fluid intelligence are statistically significant and the coefficients are identical as predictors of higher financial literacy scores. Once cognitive ability is controlled, the age coefficient remains statistically significant but the magnitude of the annual decline falls by 41%. When age is specified using 5-year age groups, none of the age groups is statistically significant after we control for word recall and vocabulary ability. Again, we find that fluid and crystallized intelligence predict financial literacy scores by a similar magnitude.

4. Conclusion

Using a new financial literacy instrument and a large, nationally representative sample, this study is the first to directly measure the decline in financial literacy in advanced age. We find a consistent linear decline in average financial literacy score of about 1 percentage point per year among respondents over 60. In order to correct for a number of potential biases due to differences in experience and incentives to invest in financial knowledge, we conduct a number of subsample analyses. We find that the effect of age on financial literacy score is strong and the magnitude of the effect is consistent across all models. The likelihood of correctly answering each of the 16 questions used in the literacy score declines significantly with age, and the rate of decline is not different within a topic area (insurance) that had higher rates of ownership among older cohorts. Confidence in financial decision-making ability does not decline with age, and the likelihood of having high confidence and a low financial literacy score rises in old age. Older respondents with higher financial literacy are more likely to make higher quality financial decisions.

In a separate analysis using the Health and Retirement Study that includes measures of cognition, we find that controlling for fluid intelligence (word recall) and crystallized intelligence (vocabulary ability) eliminates the statistical significance of age cohort categories as predictors of financial literacy and weakens the significance of a linear age variable. The results suggest that falling financial literacy scores in old age are related to the well-established decline in cognition documented in the aging literature. The results are similar to other studies of changes in financial preferences, such as reduced risk tolerance in old age, which disappear when cognitive capacities are included as predictor variables (Henninger, Madden and Huettel, 2010). The implication is that a decline in financial literacy may be a natural consequence of cognitive changes in old age.

Our results add to the literature on observed declines in financial performance with advanced age by providing evidence consistent with the decline in financial literacy being caused by a general decline in cognition. For example, Korniotis and Kumar (2011) show a decrease in investment performance that mirrors observed declines in cognitive ability by age. Our study shows that the decline in performance may be attributed directly to an age-related decrease in financial knowledge and the ability to apply knowledge correctly to financial decision-making. We are also able to better control for possible confounding effects closely related to financial knowledge by using household-level control variables such as education, homeownership, and race. For example, the proportion of individuals with a college degree declines among older cohorts, which could explain the decline in financial literacy and investment performance. However, we find that college-educated respondents over 60 experience a decline in financial literacy with age that is similar in magnitude to the full sample.

Multivariate analyses censored by financial literacy within four topic areas provide the most convincing evidence that our results are not driven by cohort effects. Life insurance ownership rates are higher among older age cohorts (Chen, Wong and Lee, 2003). Three of the four insurance questions test knowledge and application of life insurance concepts (the fourth tests knowledge of insurance deductibles). The magnitude of annual decline in scores on insurance knowledge (0.96%) is nearly identical to the decline in basic financial literacy (0.94% per year) and similar to the annual decline in investment literacy (1.02%). We find a similar annual decline among stockowners (0.98%), evidence that cohort effects related to differences in equity market participation are not driving the decline in financial literacy. Both significantly lower financial literacy scores among the youngest respondents and a consistent linear decline in financial literacy after age 60 are consistent with crystallized intelligence theory rather than a technology-based cohort effect that favors younger respondents.

Empirical evidence from cross-sectional studies of cognitive aging show a steady linear decline in tests of fluid intelligence and a more modest decline in problem solving that involves both processing ability and information retrieval. Financial literacy questions in the Health and Retirement Study (HRS) assess numerical ability (for example the estimation of compound interest over time), as well as problem solving skills (such as whether an employee should own employer stock). Our results show that, unlike the ability to solve a crossword puzzle (Salthouse, 2010), financial literacy requires both the ability to recall terms and the ability to correctly solve problems that require fluid intelligence. Fluid intelligence measured through word retrieval and crystallized intelligence measured through a vocabulary test have an equal impact on predicted financial literacy scores in the HRS.

A decline in financial skills may not lead to poor financial outcomes if individuals recognize and anticipate the decline. For example, recognition of diminished investment skills may increase demand for annuitization or the delegation of important financial decisions to a trusted advisor. Studies of trading frequency provide some evidence that older investors are less overconfident than younger investors (Barber and Odean, 2001). In contrast, our study finds that, in aggregate and within all financial decision making domains, advanced age increases overconfidence in financial decision-making abilities. The largest marginal effects are within the investment and insurance topic areas. The less educated, non-whites, and females are more likely to be financially overconfident in the old age sample. We do not find, however, that individuals who are most overconfident are less likely to optimally refinance or to choose a reward credit.

Our results show that it is not so much the imbalance between confidence and knowledge that is causing poor financial decisions, but the low financial literacy itself. Among the aged within similar decision-making domains, there is an inclination to reject evidence of declining mental abilities. For example, older drivers generally do not perceive a decline in their driving skills despite a predictable deterioration in sensory ability with advanced age (Holland and Rabbitt, 1992). However, they report that those who did perceive a decline in their abilities, and those who took an objective test that provided evidence of a decline, modify their driving behavior to reduce the likelihood of getting into an accident. It is possible that increased awareness of the natural decline in cognitive abilities essential to making effective financial decisions will lead to greater demand for more passive financial instruments such as annuities or passive investment vehicles that automatically rebalance. It may also increase demand for professional services such as financial planning, accounting and legal assistance that substitute for one's own decision-making ability. The simultaneous decline in financial literacy and increase in decision-making confidence with advanced age also has implications for national retirement policy. Programs (such as Social Security) that automatically annuitize retirement income and do not require a retiree to manage withdrawal and investment, may improve social welfare (Diamond, 2004).

References

Agarwal, S., Driscoll, J.C., Gabaix, X., and Laibson, D., 2009. The age of reason: Financial decisions over the life-cycle with implications for regulation, *Brookings Papers on Economic Activity*, **40**, 51-117.

Agarwal, S., Driscoll, J.C., Laibson, D., 2013. Optimal mortgage refinancing, *Journal of Money, Credit and Banking*, **45**(4), 591-622.

Agarwal, S., Mazumder, B., 2013. Cognitive abilities and household financial decision making, *American Economic Journal: Applied Economics*, **5**(1), 193-207.

Almenberg, J., Widmark, O., 2011. Numeracy, financial literacy, and participation in asset markets. Available at SSRN: http://ssrn.com/abstract=1756674.

Banks, J. 2010. Cognitive function, financial literacy and financial outcomes at older ages: Introduction. *The Economic Journal,* **120**(548), F357-F362.

Barber, B.M., Odean, T., 2001. Boys will be boys: Gender, overconfidence, and common stock investment, *Quarterly Journal of Economics*, **116**, 261-292.

Boyle, P.A., Yu, L., Wilson, R.S., Gamble, K., Buchman, A.S., Bennett, D.A., Reddy, H. 2012. Poor decision making is a consequence of cognitive decline among older persons without Alzheimer's disease or mild cognitive impairment, *PLoS ONE,* 7(8), e43647.

Brown, J.R., Ivkovic, Z., Smith, P.A., Weisbenner, S., 2008. Neighbors matter: Causal community effects and stock market participation, *The Journal of Finance*, **93**(3), 1509-1531.

Butrica, B.A., Iams, H.M., 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).

Bugg, J.M., DeLosh, E.L., Davalos, D.B., Davis, H.P. 2006. Age differences in fluid intelligence: Contributions of general slowing and frontal decline, *Brain and Cognition*, 62(1), 9-16

Campbell, J.Y. 2006. Household finance, *Journal of Finance*, **61**(4), 1553-1604.

Chen, R., Wong, K.A., Lee, H.C., 2003. Age, period, and cohort effects on life insurance purchases in the U.S., *The Journal of Risk and Insurance*, **68**(2), 303-328.

Ching, A.T., Hayashi, F. 2010. Payment card rewards programs and consumer payment choice, *Journal of Banking and Finance*, **34**(8), 1773-1787.

Croson, R., Gneezy, U., 2009. Gender differences in preferences, *Journal of Economic Literature*, **47**(2), 1-27.

Day, J.C., 2010. National population projections. U.S. Census Bureau, retrieved 4/24/2011 from http://www.census.gov/population/www/pop-profile/natproj.html de Bruin, W.B., Parker, A., Fischoff, B., 2012. Explaining adult age differences in decision-making competence, *Journal of Behavioral Decision Making*, **25**(4), 352-360.

DeCarli, C., Massaro, J., Harvey, D., Hald, J., Tulberg, M., Au, R., Beiser, A., D'Agostino, R. Wolf, P.A., 2005. Measures of brain morphology and infarction in the Framingham Heart Study: Establishing what is normal., *Neurobiology of Aging*, 26(4), 491-510.

Diamond, P., 2004. Social Security. *American Economic Review*, **94**(1), 1-24.

Diehl, M., Willis, S.L., Schaie, K.W., 1995. Everyday problem solving in older adults: Observational assessment and cognitive correlates, *Psychology and Aging*, 10(3), 478-491.

Gustman, A.L., Steinmeier, T.L., Tabatabai, N., 2012. Financial knowledge and financial literacy at the household level, *American Economic Review*, **102**(3), 309–313.

Henninger, D.E., Madden, D.J., Huettel, S.A. 2010. Processing speed and memory mediate age-related differences in decision making, *Psychology and Aging*, **25**(2), 262-270.

Hibbard, J.H., Slovic, P., Peters, E., Finucane, M.L., Tusler, M. 2001. Is the informed-choice policy approach appropriate for Medicare beneficiaries? *Health Affairs*, **20**(3), 199-203.

Hilgert, M.A., Hogarth, J.M., Beverly, S., 2003. Household financial management: The connection between knowledge and behavior. *Federal Reserve Bulletin* **89** (July), 309–322.

Holland, C.A., Rabbitt, P.M.A., 1992. People's awareness of their age-related sensory and cognitive deficits and the implications for road safety, *Applied Cognitive Psychology*, **6**(3), 217-231.

Horn, J.L., Cattell, R.B., 1967. Age differences in fluid and crystallized intelligence, *Acta Psychologica*, **26**, 107-129.

Hsu, J.W., Willis, R.J. 2013. Dementia risk and financial decision making by older households: The impact of information, *Journal of Human Capital*, 7(4), 340-377.

Hung, A.A., Parker, A.M., Yoong, J.K., 2009. Defining and measuring financial literacy, Rand Corporation Working Paper WR-708.

Huston, S.J., 2010. Measuring financial literacy, *Journal of Consumer Affairs*, **44**(2), 296–316.

Korniotis, G.M., Kumar, A., 2011. Do older investors make better investment decisions? *The Review of Economics and Statistics*, **93**(1), 244-265.

Lusardi, A., Mitchell, O.S., 2007a. Baby boomer retirement security: The roles of planning, financial literacy, and housing wealth, *Journal of Monetary Economics*, **54**(1), 205–224.

Lusardi, A., Mitchell, O.S., 2007b. Financial literacy and retirement planning: New evidence from the Rand American Life Panel. Michigan Retirement Research Center Working Paper 2007–157.

Lusardi, A., Tufano, P., 2009. Debt literacy, financial experiences, and overindebtedness. NBER Working Paper 14808.

Lusardi, A., Mitchell, O.S., Curto, V. 2014. Financial literacy and financial sophistication among older Americans. *Journal of Pension Economics and Finance*, forthcoming.

Malmendier, U., Nagel, S., 2011. Depression babies: Do macroeconomic experiences affect risk-taking? *Quarterly Journal of Economics*, **126**(1), 373-416.

Mandell, L. 2008. Financial literacy of high school students, In Jing Jian Xiao, ed., *Handbook of Consumer Finance Research*, 163–183. New York: Springer.

McArdle, J.J., Fisher, G.G., Kadlec, K.M. 2007. Latent variable analyses of age trends in cognition in the Health and Retirement Study, 1992-2004. *Pscyhology and Aging*, **22**(3), 525-545.

Moore, D. 2003. Survey of financial literacy in Washington state: knowledge, behavior, attitudes, and experiences. Technical Report 03-39. Washington State Department of Financial Institutions.

Peress, J., 2004. Wealth, information acquisition, and portfolio choice. *Review of Financial Studies*, **17**(3), 879-914.

Poterba, J.M., Samwick, A. 1995. Stock ownership patterns, stock market fluctuations, and consumption. *Brookings Papers on Economic Activity*, 2, 295-357.

Rushton, J.P., Ankney, C.D. 2009. Whole brain size and general mental ability: A review. *International Journal of Neuroscience*, **119**, 692-732.

Salthouse, T.A., 2009. When does age-related cognitive decline begin? *Neurobiology of Aging*, **30**(4), 507-514.

Salthouse, T.A., 2010. *Major issues in cognitive aging.* New York, NY, Oxford University Press.

Smith, J.P., McArdle, J.J., Willis, R., 2010. Financial decision making and cognition in a family context, *Economic Journal*, **120**(549), 363-380.

Stango, V., and Zinman, J. 2014. Limited and varying consumer attention: Evidence from shocks to the salience of bank overdraft fees. *The Review of Financial Studies*, **27**(4), 990-1030.

van Rooij, M., Lusardi, A., Alessie, R. 2011. Financial literacy and stock market participation, *Journal of Financial Economics*, **101**(2), 449-472.

Table 1 Univariate Relation between CFM Financial Literacy Score and Age
Each row in this table presents the results of a univariate regression of the annual change in mean financial literacy score (percentage points--pp) for each age from 60 to 94 (dependent variable) on age (independent variable)

Financial Literacy Measure	*Annual Change (pp)*	*T-statistic*	*R-Square*
Overall Financial Literacy Score (16 Questions, %)	-1.53	26.27***	0.96
Overall Confidence in Literacy	0.11	1.86	0.07
Overconfidence in literacy	1.63	19.88***	0.92
Basic Literacy	-1.42	20.86***	0.93
Borrowing Literacy	-1.61	22.30***	0.94
Investment Literacy	-1.65	21.48***	0.94
Insurance Literacy	-1.43	18.75***	0.92
Confidence in Managing Money	0.06	1.12	0.01
Confidence in Credit	0.05	0.99	0.00
Confidence in Investing	-0.00	0.05	0.00
Confidence in Insurance	0.31	3.91***	0.31

***, **,* indicate significance at the 0.01, 0.05 and 0.10 levels, respectively

Table 2a **CFM Sample Financial Literacy Scores and Variable Frequency/Mean**

Each panel in this table presents the total financial literacy score and topic area financial literacy scores for each variable, along with the sample frequency or mean for each variable attribute.

	Financial Literacy Objective Scores (0-100)						
	Total Score %/(std)	*Basics %/(std)*	*Borrowing %/(std)*	*Investment %/(std)*	*Insurance %/(std)*	*Frequency Or Mean*	*N*
Age							
Age 60-69	62 (26)	62 (31)	63 (30)	57 (34)	65 (31)	51.1%	1979
Age 70-79	49 (28)	48 (33)	51 (32)	43 (36)	54 (34)	31.0%	1201
Age 80 or above	32 (24)	34 (29)	33 (29)	25 (29)	37 (32)	17.9%	693
Education							
<High School	25 (20)	22 (25)	30 (28)	16 (22)	31 (30)	4.9%	190
High School	39 (26)	38 (30)	42 (32)	31 (32)	44 (33)	23.8%	922
Some College	51 (26)	51 (32)	54 (32)	44 (34)	55 (32)	25.9%	1003
College	61 (26)	62 (31)	61 (31)	58 (34)	64 (31)	24.5%	949
Graduate	66 (26)	67 (30)	64 (30)	62 (34)	69 (31)	20.9%	809
Race							
White	54 (28)	54 (33)	55 (32)	48 (36)	58 (33)	88.6%	3431
Non-White	42 (27)	40 (31)	47 (33)	35 (33)	47 (34)	11.4%	442
Gender							
Male	58 (28)	58 (33)	59 (32)	54 (36)	61 (33)	43.3%	1677
Female	48 (28)	48 (33)	50 (32)	41 (34)	53 (34)	56.7%	2196
Homeownership							
Homeowner	54 (28)	54 (33)	55 (32)	49 (36)	58 (33)	90.2%	3493
Not Homeowner	38 (28)	40 (33)	40 (32)	30 (32)	43 (33)	9.8%	380
Marital Status							
Married	59 (27)	59 (32)	60 (32)	54 (35)	62 (33)	54.6%	2115
Unmarried	45 (28)	44 (32)	47 (32)	38 (34)	50 (34)	45.4%	1758
Tax Sheltered Status							
Tax Sheltered [illegible]	64 (26)	66 (30)	63 (31)	62 (34)	67 (30)	30.1%	1166
No Tax Sheltered	48 (28)	48 (33)	51 (32)	41 (35)	53 (34)	69.9%	2707
Stock Ownership							
Stock/MF Owner	64 (25)	64 (31)	63 (30)	61 (33)	67 (31)	43.1%	1669
No Stock/MF	45 (28)	45 (32)	48 (32)	37 (34)	50 (33)	56.9%	2204
Income Level							
Lowest Income	39 (27)	37 (31)	42 (32)	32 (33)	46 (33)	$7,953	775
Quintile 2 Income	46 (27)	46 (32)	48 (32)	37 (34)	51 (33)	18,881	775
Quintile 3 Income	57 (25)	55 (31)	59 (29)	51 (34)	61 (31)	33,756	775
Quintile 4 Income	62 (24)	64 (30)	63 (29)	56 (33)	65 (31)	59,427	774

Highest Income	69 (22)	70 (27)	69 (28)	66 (32)	72 (27)	165,598	774
Wealth Level							
Lowest Net Worth	37 (27)	37 (32)	40 (33)	30 (32)	42 (34)	($11,418)	775
Quintile 2 NW	44 (27)	43 (32)	48 (32)	34 (33)	50 (33)	104,917	775
Quintile 3 NW	53 (27)	54 (33)	55 (32)	47 (35)	57 (33)	250,351	775
Quintile 4 NW	60 (25)	61 (29)	60 (30)	56 (33)	64 (31)	459,537	774
Highest Net Worth	68 (24)	68 (29)	67 (29)	67 (32)	71 (29)	1,797,815	774

Table 2b CFM Sample Financial Confidence Score and Overconfidence Status
Each panel in this table presents the financial confidence score and the percentage that are overconfident for each variable attribute.

	Financial Confidence Scores (0-100) and Overconfidence Status		
	Confidence Score %/(std)	*Overconfidence %*	*N*
Age			
Age 60-69	72 (19)	5.7%	1979
Age 70-79	75 (19)	10.4%	1201
Age 80 or above	74 (21)	19.3%	693
Education			
<High School	69 (25)	16.2%	190
High School	73 (21)	16.6%	922
Some College	72 (20)	8.7%	1003
College	75 (17)	6.5%	949
Graduate	74 (18)	4.9%	809
Race			
White	73 (19)	9.0%	3431
Non-White	70 (23)	14.2%	442
Gender			
Male	74 (19)	7.2%	1677
Female	73 (19)	11.5%	2196
Homeownership			
Homeowner	74 (19)	9.4%	3493
Not Homeowner	66 (22)	11.9%	380
Marital Status			
Married	74 (18)	7.4%	2115
Unmarried	72 (21)	12.2%	1758
Tax Sheltered Status			
Tax Sheltered Acct.	76 (16)	3.8%	1166
No Tax Sheltered	72 (20)	11.1%	2707

Stock Ownership			
Stock/MF Owner	77 (16)	5.0%	1669
No Stock/MF	72 (20)	10.7%	2204
Income Level			
Lowest Income	69 (23)	13.1%	775
Quintile 2 Income	70 (22)	13.7%	775
Quintile 3 Income	71 (19)	7.8%	775
Quintile 4 Income	74 (17)	5.2%	774
Highest Income	76 (17)	2.9%	774
Wealth Level			
Lowest Net Worth	65 (23)	11.5%	775
Quintile 2 NW	69 (21)	11.7%	775
Quintile 3 NW	73 (18)	8.4%	775
Quintile 4 NW	75 (17)	5.9%	774
Highest Net Worth	79 (15)	4.4%	774

Table 3 CFM Regressions for Financial Literacy Scores and Age

The first column presents the results of a univariate regression financial literacy score (% correct out of 16 questions – dependent variable) on continuous age (independent variable). The second column presents the same regression using 5-year age categories. The third (age continuous) and fourth (age categorical) columns present multivariate regressions that include a matrix of control variables.

	Linear Age	*Cohort Age*	*Linear Age*	*Cohort Age*
Age	-1.36***		-1.02***	
	(0.05)		(0.05)	
Age 65-69		-1.62		-1.41
		(1.14)		(1.01)
Age 70-74		-9.85***		-5.98***
		(1.22)		(1.10)
Age 75-79		-17.36***		-11.8***
		(1.33)		(1.22)
Age 80-84		-26.66***		-19.95***
		(1.44)		(1.31)
Age 85-89		-32.55***		-25.41***
		(1.87)		(1.71)
Age 90+		-41.72***		-32.15***
		(3.26)		(3.14)
<High School			-10.09***	-10.23***
			(1.79)	(1.79)
Some College			7.92***	8.13***
			(1.03)	(1.03)
College			13.85***	14.11***
			(1.08)	(1.08)
Graduate			16.83***	16.93***
			(1.14)	(1.14)
High Income			2.38*	2.72**
			(1.20)	(1.19)
High Wealth			5.47***	5.08***
			(0.99)	(0.99)
White			8.82***	8.86***
			(1.15)	(1.15)
Male			4.35***	4.28***
			(0.76)	(0.76)
Homeowner			4.37***	4.08***
			(1.26)	(1.26)
Married			3.96***	3.76***
			(0.80)	(0.80)
Tax Sheltered			5.42***	5.32***
			(0.87)	(0.87)
Stock/MF			5.43***	5.77***
			(0.84)	(0.84)
Sample Size	4152	4152	3898	3898
Adjusted R^2	0.17	0.17	0.37	0.37

***,**,* indicate significance at the 0.01, 0.05 and 0.10 levels, respectively

Table 4 CFM Censored Regressions of Financial Literacy and Age
Each column of this table presents the results of a multivariate regression of financial literacy (dependent variable) on age (independent variable) and control variables. The sample is censored by those with a college degree (first column), males only (second column), stockowners (third column), and the older cohort of respondents 60 and older (fourth column).

	College	*Males*	*Stock/MF Owners*	*Older Cohort*
Age	-1.1***	-0.96***	-0.98***	-1.37***
	(0.06)	(0.07)	(0.08)	(0.22)
<High School		-11.19***	-6.79	-5.53
		(3.25)	(4.89)	(3.39)
Some College		8.52***	7.54***	5.00**
		(1.72)	(1.81)	(2.12)
College		15.95***	11.89***	11.04***
		(1.73)	(1.77)	(2.31)
Graduate		17.22***	15.37***	12.7***
		(1.75)	(1.78)	(2.54)
High Income	3.53**	2.25	2.64	4.29
	(1.34)	(1.68)	(1.62)	(4.13)
High Wealth	5.3***	4.84***	5.35***	0.77
	(1.09)	(1.35)	(1.27)	(2.43)
White	8.75***	8.36***	6.82***	5.47**
	(1.42)	(1.75)	(1.89)	(2.56)
Male	5.4***		4.15***	6.76***
	(0.9)		(1.15)	(1.72)
Homeowner	4.67***	6.84***	1.16	-1.14
	(1.67)	(2.21)	(2.82)	(2.25)
Married	4.37***	6.53***	3.59***	2.44
	(0.97)	(1.26)	(1.26)	(1.79)
Tax Sheltered	6.04***	6.16***	4.04***	9.26***
	(0.98)	(1.27)	(1.16)	(1.97)
Stock/MF	5.25***	4.08***		7.36***
	(0.98)	(1.26)		(1.93)
Sample Size	2,803	1702	1,654	800
Adjusted R^2	0.27	0.34	0.23	0.26

***,**,* indicate significance at the 0.01, 0.05 and 0.10 levels, respectively

Table 5 **CFM Regressions of Financial Literacy Topic Areas and Age**

Each column of this table presents the results of a multivariate regression of individual financial literacy topic areas (dependent variables) on age (dependent variable) and control variables. The first column dependent variable is the four questions on financial literacy basic topics, the second column dependent variable is the four questions on borrowing, the third column dependent variable is the four financial literacy questions on investing, and the fourth column dependent variable is the four insurance-related financial literacy questions.

	Basics	**Borrowing**	**Investment**	**Insurance**
Age	-0.94***	-1.1***	-1.06***	-0.96***
	(0.06)	(0.06)	(0.06)	(0.06)
<High School	-11.87***	-9.76***	-9.95***	-8.8***
	(2.23)	(2.27)	(2.36)	(2.37)
Some College	8.55***	7.07***	8.17***	7.88***
	(1.28)	(1.31)	(1.36)	(1.37)
College	15.9***	10.79***	15.72***	12.99***
	(1.34)	(1.37)	(1.42)	(1.43)
Graduate	18.78***	13.27***	19.08***	16.18***
	(1.42)	(1.45)	(1.5)	(1.51)
High Income	2.76*	1.55	3.16**	2.05
	(1.53)	(1.49)	(1.46)	(1.41)
High Wealth	4.56***	3.94***	8.44***	4.94***
	(1.23)	(1.28)	(1.29)	(1.31)
White	11.2***	6.47***	9.99***	7.61***
	(1.42)	(1.46)	(1.51)	(1.52)
Male	4.28***	3.95***	6.23***	2.96***
	(0.94)	(0.96)	(0.99)	(1.00)
Homeowner	1.1	6.17***	4.82***	5.4***
	(1.57)	(1.6)	(1.66)	(1.67)
Married	4.62***	4.11***	3.65***	3.47***
	(0.99)	(1.01)	(1.05)	(1.05)
Tax Sheltered	6.81***	3.32***	7.29***	4.25***
	(1.08)	(1.1)	(1.14)	(1.15)
Stock/MF	5.89***	3.15***	7.82***	4.87***
	(1.03)	(1.05)	(1.1)	(1.11)
Sample Size	3,898	3,898	3,898	3,898
Adjusted R^2	0.29	0.23	0.32	0.22

***,**,* indicate significance at the 0.01, 0.05 and 0.10 levels, respectively

Table 6 **CFM Logistic Regression Age Coefficients for Individual Financial Literacy Items**
This table presents the age coefficients (only) from 16 logistic regressions of each individual financial literacy item (dependent variable) on age (independent variable) and household control variables (education, income, wealth levels, race, gender, marital status and ownership of home, tax sheltered account, and stocks).

Financial Literacy Question	*Age Pt. Estimate*	*St. Beta*	*Max-rescaled R^2*
Net worth is equal to:	-5.2%***	0.23[c]	0.26
If your assets increase by $5,000 and your liabilities decrease by $3,000, your net worth would	-5.0%***	0.2 [a]	0.18
Which bank account is likely to pay the highest interest rate on money saved?	-6.2%***	0.28[a]	0.18
Savings accounts and money market accounts are most appropriate for:	-7.0%***	0.32[a]	0.20
To reduce the total finance costs paid over the life of an auto loan, you should choose a loan with the	-5.7%***	0.26[c]	0.28
If you always pay the full balance on your credit card, which of the following is least important?	-6.4%***	0.29[a]	0.21
On which type of loan is interest never tax deductible?	-7.0%***	0.32[a]	0.25
Which type of mortgage would allow a first-time home buyer to qualify for the highest loan amount?	-2.9%***	0.13[c]	0.15
The benefit of owning investments that are diversified is that it	-6.0%***	0.27[a]	0.19
A young investor willing to take moderate risk for above-average growth would be most interested in:	-4.2%***	0.19[a]	0.19
The main advantage of a 401(k) plan is that it:	-4.3%***	0.20[a]	0.12
To ensure that some of your retirement savings will not be subject to income tax upon withdrawal, you would contribute to:	-3.8%***	0.17[a]	0.11
If you have an insurance policy with a higher deductible, the premiums will be:	-3.9%***	0.17[b]	0.17
Which of the following types of insurance is most important for single workers without children?	-5.4%***	0.24[a]	0.22
Which policy provides the most coverage at the lowest cost for a young family?	-4.7%***	0.21[a]	0.12
Which household would typically have the greatest life insurance need?	-3.9%***	0.17[b]	0.11

*** indicates significance at the 0.0001 levels.
[a,b,c,d] indicates age ranks 1st, 2nd, 3rd, &4th, respectively, of explaining variation in answering question correctly.

Table 7 **CFM Financial Confidence Regressions on Age**

Each column in this table presents multivariate regression results of financial confidence on age and household control variables. The first column dependent variable is total confidence that combines all of the four confidence items. The second through fifth columns use each of the four confidence items individually as the dependent variable (respondent-reported confidence in managing money, managing credit, using investment products and using insurance products, respectively). The last column presents the logistic regression results of overconfidence (yes/no) as the dependent variable on age (independent variable) along with the matrix of control variables.

	Total Confidence	*Managing Money*	*Managing Credit*	*Using Investments*	*Using Insurance*	*Over-Confident*
Age	0.03***	0.01***	0.02***	0.04***	0.04***	1.07***
	(0.00)	(0)	(0)	(0.01)	(0.01)	
Objective Score (by area)	0.01***	0.00**	0.00**	0.01***	0.00	
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	
<High School	-0.16	0.06	-0.33**	-0.14	-0.24	1.48
	(0.18)	(0.16)	(0.17)	(0.27)	(0.26)	
Some Coll.	-0.11	-0.08	-0.14	0.03	-0.18	0.59***
	(0.09)	(0.09)	(0.09)	(0.14)	(0.14)	
College	-0.06	-0.09	-0.05	0.2	-0.13	0.45***
	(0.10)	(0.1)	(0.1)	(0.15)	(0.15)	
Graduate	-0.27***	-0.27***	-0.22**	-0.12	-0.32**	0.35***
	(0.10)	(0.1)	(0.1)	(0.16)	(0.16)	
High Income	0.21**	0.12	0.14	0.27*	0.21	0.69
	(0.09)	(0.09)	(0.09)	(0.14)	(0.14)	
High Wealth	0.35***	0.26***	0.42***	0.64***	0.18	0.79
	(0.09)	(0.09)	(0.09)	(0.14)	(0.13)	
White	0.17*	0.16	0.38***	0.02	0.13	0.68**
	(0.10)	(0.1)	(0.1)	(0.16)	(0.16)	
Male	-0.03	-0.13**	-0.15**	0.28***	-0.13	0.73**
	(0.07)	(0.07)	(0.07)	(0.1)	(0.1)	
Homeowner	0.47***	0.19*	0.36***	0.9***	0.36**	1.3
	(0.12)	(0.11)	(0.12)	(0.18)	(0.17)	
Married	-0.02	-0.02	-0.02	-0.03	0.06	0.9
	(0.07)	(0.07)	(0.07)	(0.11)	(0.11)	
Tax Sheltered	0.31***	-0.03	0.2***	0.87***	0.2*	0.94
	(0.07)	(0.08)	(0.08)	(0.12)	(0.12)	
Stock/MF	0.38***	0.16**	0.33***	0.63***	0.29**	0.93
	(0.07)	(0.07)	(0.07)	(0.11)	(0.11)	
Sample Size	3,455	3,849	3,815	3,590	3,657	3,455
Adjusted R^2	0.07	0.01	0.04	0.10	0.01	0.16

***,**,* indicate significance at the 0.01, 0.05 and 0.10 levels, respectively

Table 8 **CFM Regressions for Decision Quality on Financial Literacy, Confidence and Age**
Each column in this table presents multivariate regression results of decision quality outcomes (dependent variable) on age, financial literacy, financial confidence and age interacted with financial literacy and financial confidence, controlling for household characteristics. The logistic regression dependent variables are having a low mortgage rate (first column), taking advantage of reward cards for convenience credit card users (second column), and the interviewer's assessment of the respondent having a good understanding of the survey questions (third column).

	Low Mortgage Rate	*Reward Card*	*Good Understanding*
Age (60-69 reference)			
70-79	0.76	1.86**	1.03
80 and older	0.9	1.84*	1.02
Financial Literacy Score (FinLit)	1.01*	1.02***	1.06***
Overconfident (OC)	0.97	1.58	1.15
*FLS*70-79*	1.01	1	0.99
*FLS*80+*	1	1.01	0.98*
*OC*70-79*	0.87	0.92	1.25
*OC*80+*	Abnormal value	1.03	0.56
<High School	0.44	0.32***	0.78
Some College	0.92	0.99	1.17
College	2.08**	1.37**	1.08
Graduate	1.08	1.59***	1.33
High Income	1.84**	1.05	0.84
High Wealth	2.59***	2.39***	1
White	1.22	1.68***	1.79***
Male	1	0.79**	0.97
Married	1.47*	1.35***	0.9
Sample Size	998	2,511	3,640
R^2	0.20	0.30	0.32

***,**,* indicate significance at the 0.01, 0.05 and 0.10 levels, respectively.

Table 9 **HRS Regressions of Financial Literacy Scores, Age, and Cognitive Ability**
The first column presents the results of a simple univariate regression financial literacy score (% correct out of 7 HRS financial literacy-related questions – dependent variable) on continuous age (independent variable). The second column presents the same regression using age categories. The third (age continuous) and fourth (age categorical) columns present multivariate regressions that include cognitive ability measures (Word Recall and Vocabulary Ability) along control variables.

	Linear Age	*Cohort Age*	*Linear Age*	*Cohort Age*	*Linear Age*	*Cohort Age*
Age	-0.41***		-0.35***		-0.24**	
	(0.08)		(0.08)		(0.1)	
Age 65-69		0.31		1.08		1.89
		(2.14)		(1.92)		(2.24)
Age 70-74		-2.4		-0.98		-0.21
		(1.84)		(1.65)		(1.94)
Age 75-79		-4.51**		-2.69		-1.28
		(2.04)		(1.85)		(2.12)
Age 80-84		-5.71**		-6.32***		-3.64
		(2.67)		(2.39)		(2.63)
Age 85-89		-8.21***		-6.7**		-2.91
		(2.93)		(2.64)		(2.91)
Age 90+		-12.61***		-6.27*		-2.93
		(4.13)		(3.75)		(3.98)
Word Recall					0.13***	0.14***
					(0.04)	(0.05)
Vocabulary Ability					0.13***	0.13***
					(0.04)	(0.04)
<High School			-8.59***	-8.58***	-6.36***	-6.37***
			(1.72)	(1.73)	(1.93)	(1.95)
Some College			2.6*	2.9*	0.88	0.89
			(1.56)	(1.57)	(1.79)	(1.8)
College			10.52***	10.98***	7.4***	7.42***
			(1.69)	(1.71)	(2.01)	(2.03)
High Income			0.81	1.37	0.85	0.86
			(1.67)	(1.67)	(2.01)	(2.03)
High Wealth			0.27	0.08	0.95	0.91
			(1.74)	(1.76)	(1.99)	(2)
White			3.75**	3.19**	1.44	0.85
			(1.59)	(1.59)	(1.86)	(1.83)
Male			5.74***	5.58***	7.21***	7.3***
			(1.22)	(1.23)	(1.42)	(1.44)
Homeowner			3.91**	3.72**	3.1*	3.05*
			(1.59)	(1.6)	(1.8)	(1.81)
Married			-0.12	0.14	-0.22	-0.08
			(1.4)	(1.41)	(1.6)	(1.61)
Tax Sheltered			6.42***	6.46***	5.69***	5.79***
			(1.39)	(1.4)	(1.58)	(1.6)
Stock/MF			3.25**	2.98*	3.46**	3.4**
			(1.53)	(1.54)	(1.71)	(1.72)
Sample Size	1109	1109	1108	1108	887	887
Adjusted R^2	0.02	0.01	0.23	0.22	0.23	0.23

Figure 1 **Mean and Predicted Financial Literacy Score by Age**
The figure shows mean CFM financial literacy scores for each year of age and predicted financial literacy score using regression estimates when age is specified as age and age-squared, and as 5-year age groups, in a regression model that controls for household characteristics.

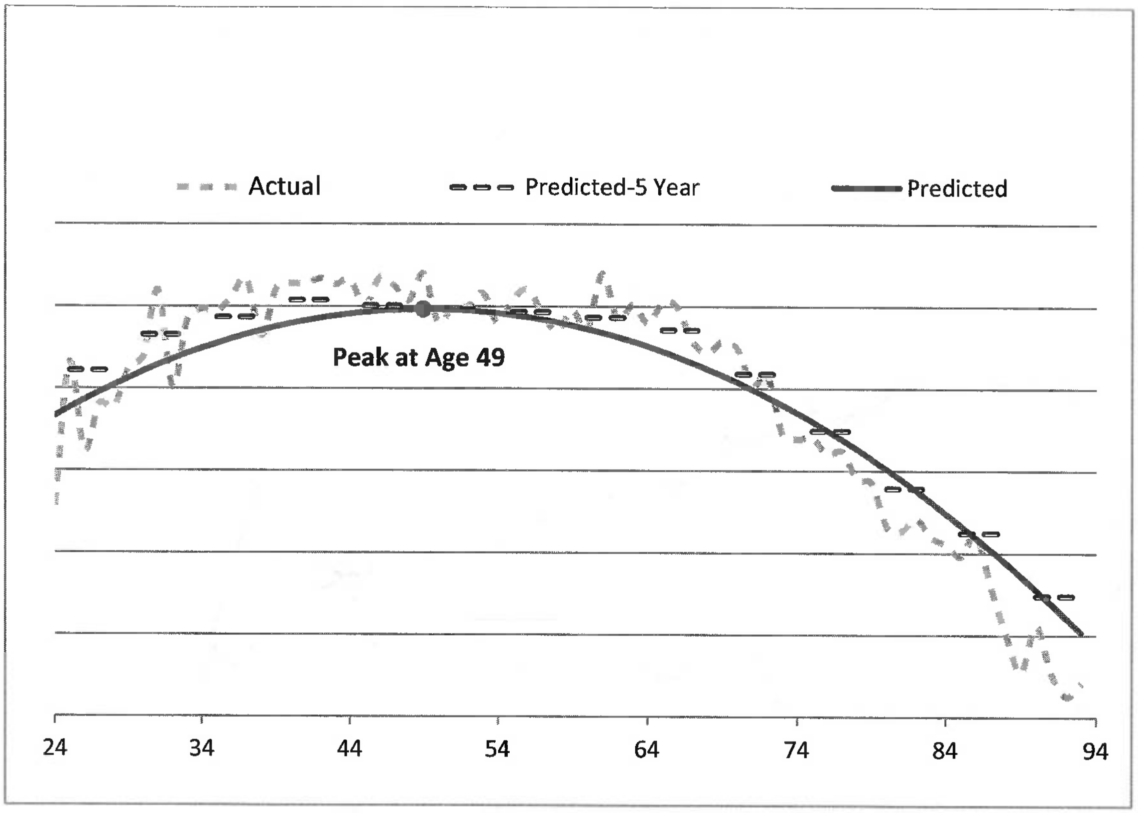


Figure 2 **Financial Literacy, Financial Confidence and Cognitive Ability**

The figure shows average financial literacy score, average confidence in financial decision making ability, and average percentage of overconfident respondents within each age year using the CFM. Average word recall score within each year of age is drawn from the HRS.

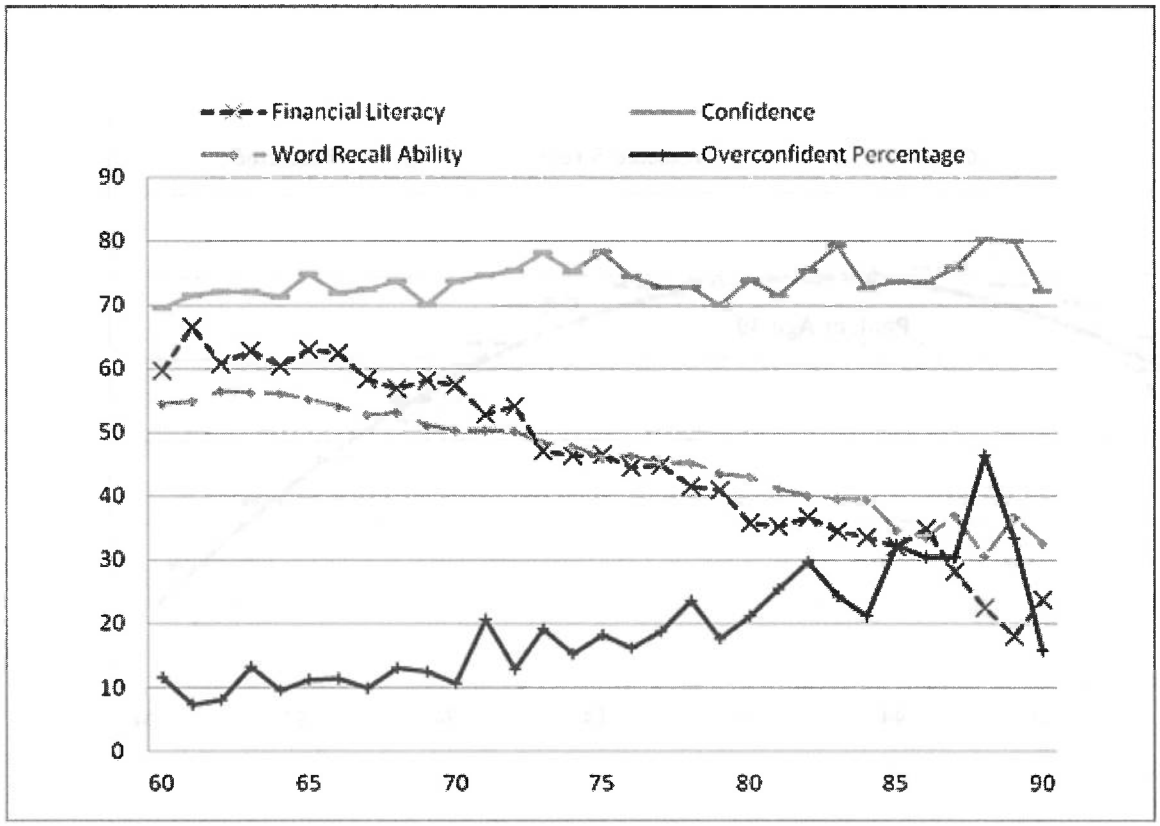
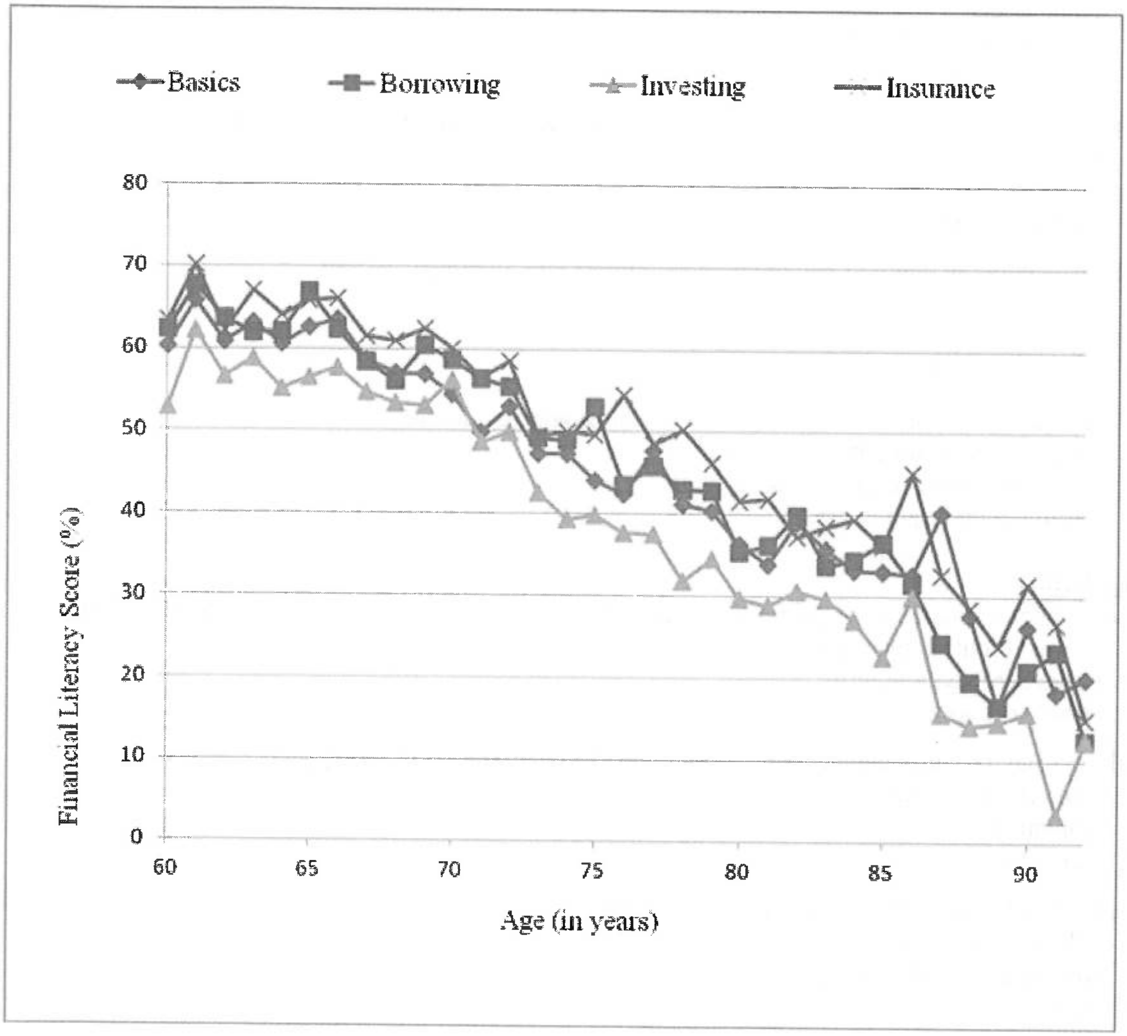


Figure 3 **Financial Literacy and Topic Area**

The figure shows average financial literacy score within each topic area (basics, borrowing, insurance and investments) by each year of age using the CFM.



Appendix 1-A Financial Literacy Assessment Test (FLAT) Items from the CFM Survey.

Basics Items:

1. Net worth is equal to:
    1. Total assets
    2. *Total assets plus liabilities*
    3. Total assets minus liabilities

2. If your assets increase by $5,000 and your liabilities decrease by $3,000, your net worth would
    1. Increase by $2,000
    2. *Increase by $8,000*
    3. Increase by $3,000

3. Which bank account is likely to pay the highest interest rate on money saved?
    1. Savings account
    2. Six month CD or certificate of deposit
    3. *Three year CD*

4. Savings accounts and money market accounts are most appropriate for
    1. Long-term investments like retirement
    2. *Emergency funds and short-term goals*
    3. Earning a high rate of return

Borrowing Items:

5. To reduce the total finance costs paid over the life of an auto loan, you should choose a loan with the
    1. Lowest monthly payment
    2. Longest repayment term
    3. *Shortest repayment*

6. If you always pay the full balance on your credit card, which of the following is least important?
    1. *Annual interest rate*
    2. Annual fees
    3. Line of credit

7. On which type of loan is interest never tax deductible?
    1. A home equity loan
    2. An adjustable rate mortgage
    3. *A personal vehicle loan*

8. Which type of mortgage would allow a first-time home buyer to qualify for the highest loan amount?
    1. Fixed-rate mortgage
    2. *Adjustable-rate mortgage*
    3. Reverse mortgage

Investing Items:

9. The benefit of owning investments that are diversified is that it
    1. *Reduces risk*
    2. Increases return
    3. Reduces tax liability

10. A young investor willing to take moderate risk for above-average growth would be most interested in:
    1. Treasury bills
    2. Money market mutual funds
    3. *Balanced stock funds*

11. The main advantage of a 401(k) plan is that it:
    1. Provides a high rate of return with little risk
    2. *Allows you to shelter retirement savings from taxation*
    3. Provides a well-diversified mix of investment assets

12. To ensure that some of your retirement savings will not be subject to income tax upon withdrawal, you would contribute to:
    1. A Traditional IRA or Individual Retirement Account
    2. *A Roth IRA*
    3. A 401(k) plan

Insurance Items:

13. If you have an insurance policy with a higher deductible, the premiums will be
    1. Higher
    2. *Lower*
    3. The same

14. Which of the following types of insurance is most important for single workers without children?
    1. Life insurance
    2. *Disability income insurance*
    3. Dental insurance

15. Which policy provides the most coverage at the lowest cost for a young family?
    1. *Renewable term life*
    2. Whole life
    3. Universal life

16. Which household would typically have the greatest life insurance needs?
    1. A middle-class retired couple
    2. A middle-aged working couple with children in college
    3. *A single-earner family with two young children in pre-school*

Confidence Items:

For the following 4 questions, record a number from 0 (LOWEST confidence) to 10 (HIGHEST confidence):

LOWEST CONFIDENCE
HIGHEST CONFIDENCE

|----------|----------|----------|----------|----------|----------|----------|----------|----------|----------|
0 1 2 3 4 5 6 7 8 9
10

1. How confident are you with managing money? _____
2. How confident are you with managing credit and debt? _____
3. How confident are you with using investment products? _____
4. How confident are you with using insurance products? _____

Appendix 1-B Financial Literacy-related Items from the HRS Survey.

Question 1.

First, suppose you had $100 in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow -- more than $102, exactly $102, or less than $102?

1. *More than $102*
2. Exactly $102
3. Less than $102

Question 2.

Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After 1 year, would you be able to buy more than today, exactly the same as today, or less than today with the money in this account?

1. More than today
2. Exactly the same as today
3. *Less than today*

Question 3.

Do you think that the following statement is true or false: buying a single company stock usually provides a safer return than a stock mutual fund?

1 True

5 *False*

Question 4.

Which asset do you think historically has paid the highest returns over a long time period, say 20 years or more -- savings accounts, bonds, or stocks?

1. Saving accounts
2. Bonds
3. *Stocks*

Question 5.

An employee of a company with publicly traded stock should have a lot of his or her retirement savings in the company's stock.

1 True

5 *False*

Question 6.

It is best to avoid owning stocks of foreign companies.

1 True

5 *False*

Question 7.

If the interest rate falls, bond prices will rise.

1 *True*

5 False

40

The Impact of the Broker-Dealer Fiduciary Standard on Financial Advice

Michael Finke
Texas Tech University

Thomas Langdon
Roger Williams University

March 9, 2012

Abstract: Consumers who rely on the financial advice of experts are at an information disadvantage that may be exploited by advisers who are not required to make recommendations that are in the best interest of the customer. Registered representatives of broker-dealers are subject to a suitability standard under the Securities Exchange Act of 1934, while investment advisers are regulated as fiduciaries under the Investment Advisers Act of 1940. An early legislative version of the 2010 Dodd-Frank Act would have eliminated the broker-dealer exception from the definition of investment adviser under the Advisers Act. If enacted, this change would have subjected brokers to a common-law fiduciary standard (like investment advisers), but was postponed to examine the consequences of this policy change. It has been suggested that the imposition of a fiduciary standard on registered representatives would result in significant changes in how broker-dealers conduct business by limiting a representative's ability to recommend commission investments, provide advice to middle-market clients, and offer a broad range of financial products. We take advantage of differences in state broker-dealer common law standards of care to test whether a relatively stricter fiduciary standard of care impacts the ability to provide services to consumers. We find that the number of registered representatives doing business within a state as a percentage of total households does not vary significantly among states with stricter fiduciary standards. A sample of advisers in states that have either a strict fiduciary standard or no fiduciary standard are asked whether they are constrained in their ability to recommend products or serve lower-wealth clients. We find no statistical differences between the two groups in the percentage of lower-income and high-wealth clients, the ability to provide a broad range of products including those that provide commission compensation, the ability to provide tailored advice, and the cost of compliance.

Keywords: Fiduciary regulation, broker dealer exemption, financial advice, household finance, investment advising, brokerage industry

---

This research was made possible by donations from the Roger and Brenda Gibson Family Foundation, fi360, the Committee for the Fiduciary Standard, and the Financial Planning Association.

I. Introduction

Financial advisers provide expert assistance selecting financial instruments for retail customers. Registered representatives of broker-dealers facilitate the sale of securities and often provide financial advice to clients who are less knowledgeable about the product. This imbalance of information has led to the imposition of a legal fiduciary standard when an informed agent is hired to make decisions on behalf of a less informed client (Frankel, 1983). In the absence of an informational imbalance between registered representatives (or brokers) and their customers, the primary service provided through broker-dealers is to sell retail financial products demanded by the customer. However, many broker-dealers have suggested through advertising and by referring to registered representatives with terms such as "financial planner" or "financial consultant" that their services include planning or consulting services that involve the provision of expert advice (Hung, Clancy, Dominitz, Talley, Berrebi and Suvankulov, 2008). Most consumers assume that advising services are provided by registered representatives of broker-dealers (Hung et al., 2008).

While consumers are generally unable to distinguish between investment advisers whose primary purpose is to provide investment advice and registered representatives whose advice is considered incidental to the sale of financial products, they are regulated by two different entities that apply different market conduct standards. Investment advisers are regulated by the Securities and Exchange Commission (SEC or Commission) under the Investment Advisers Act of 1940 (Advisers Act) as fiduciaries and a fiduciary standard of care is applied to the advice given to their clients. Registered representatives of broker-dealers are regulated under the Securities Exchange Act of 1934 through the Financial Industry Regulatory Authority (FINRA), a self-regulatory organization. Registered representatives must meet a standard of suitability when providing information about financial products, and are not assumed to have a fiduciary responsibility toward customers.

The difference in regulation between investment advisers and brokers impacts the market for financial advice. The sale of professional advisory services to a less-informed client involves significant potential agency costs that exist when the interests of the client and broker/adviser are not perfectly aligned (Jensen and Meckling, 1976). These costs occur when the broker recommends products that benefit the broker to the disadvantage of the customer. Examples of agency costs include recommending products that have higher commissions or not taking the time to consider alternative financial strategies for a customer. It is possible that the application of a suitability standard to investment advice will lead to greater agency costs. A suitability constraint allows brokers to recommend products that are not necessarily in the best interest of the client but may be considered potentially suitable given the customer's characteristics and needs. This latitude in product recommendation among registered representatives provides a greater opportunity to extract customer rents than would be possible under the constraints of a fiduciary standard (Cummings and Finke, 2010). If the suitability standard provides greater opportunities to extract rents from clients, we would expect the broker-dealer industry to defend its ability to maintain this advantage by continuing the existing regulatory regime.

If, however, a fiduciary standard was applied to registered representatives whose sole purpose is to facilitate the sale of financial instruments within a competitive marketplace, the imposition of a fiduciary standard to these sales activities may have a negative impact on the ability of broker-dealers to provide a variety of financial products to consumers. Many consumers may demand products whose appropriate use is difficult for a registered representative to defend as being in the customer's best interest. For example, there may be mutual funds that pay a commission to the broker that are less efficient than comparable mutual funds that pay no commission. The brokerage industry has argued that since moderate income clients are less attractive to investment advisers, who are often compensated based on a percentage of assets under management, these clients often seek financial advice from registered representatives compensated through product commissions (Headley, 2011). These less wealthy clients may be less able to receive much-needed financial advice incidental to the sale of commission products if brokers incur increased liability under a fiduciary standard. The application of a standard of care that assumes a fiduciary relationship between registered representative and customer may constrain the ability to make product recommendations and limit the range of available financial products.

While the industry has suggested that fiduciary regulation will have an adverse impact on the industry, there are no existing empirical studies that examine the impact of a change in regulatory policy on the marketplace for financial advice. This study takes advantage of heterogeneity in broker-dealer regulation among states to test whether a relatively more strict application of a common law fiduciary standard of care impacts the number of registered representatives doing business within the state. We also conduct a survey to assess differences in perceived ability to provide financial products among states subject to stricter fiduciary standards. We find that the saturation of registered representatives within states does not vary significantly among states with different fiduciary regulation. When advisers in states that have a stricter fiduciary standard are asked whether they are constrained in their ability to recommend products, or if they are unable to serve lower-wealth clients, we find no statistical difference between advisers from states that do and do not apply a common law fiduciary standard.

II. Background

On July 15, 2010, Congress passed the Dodd-Frank Wall Street Reform and Consumer Protection Act (Dodd-Frank Act). Section 913 of the Dodd-Frank Act required the SEC to conduct a study to evaluate, among other things, (1) the effectiveness of existing legal or regulatory standards of care (imposed by the Commission, a national securities association, and other federal or state authorities) for providing personalized investment advice and recommendations about securities to retail customers; and (2) whether there are legal or regulatory gaps, shortcomings, or overlaps in legal or regulatory standards in the protection of retail customers relating to the standards of care for providing personalized investment advice about securities to retail customers that should be addressed by rule or statute. In one of the early legislative drafts, Dodd-Frank would have eliminated the broker-dealer exception from the definition of investment adviser under the Advisers Act, but the legislation as adopted included a compromise to conduct further study of the issue. The Dodd-Frank Act defines "retail customer" as a natural person, or the legal representative of a natural person, who receives personalized investment advice about securities from a broker or dealer or investment adviser and who uses that advice for personal, family, or household purposes.

In January 2011, the SEC released its Study on Investment Advisers and Broker-Dealers (Staff of the U.S. Securities and Exchange Commission, 2011). In its report, the SEC staff noted that "the regulatory regime that governs the provision of investment advice to retail investors is essential to assuring the integrity of that advice and to matching legal obligations with the expectations and needs of investors," and found that investors are often confused by differing standards of care that apply to investment advisers and broker-dealers (Staff of the U.S. Securities and Exchange Commission, 2011). The SEC study recommended the adoption of a uniform fiduciary standard for investment advisers and broker-dealers that provides:

> The standard of conduct for all brokers, dealers, and investment advisers, when providing personalized investment advice about securities to retail customers (and such other customers as the Commission may by rule provide), shall be to act in the best interest of the consumer without regard to the financial or other interest of the broker, dealer, or investment adviser providing the advice (Staff of the U.S. Securities and Exchange Commission, 2011).

The SEC study recommends that the Commission, in implementing a uniform fiduciary standard, should engage in rulemaking and provide interpretive guidance addressing the two major components of a uniform fiduciary standard: the duties of loyalty and care. When addressing the duty of loyalty, the report suggests that a uniform fiduciary standard will obligate both investment advisers and broker-dealers to eliminate or disclose conflicts of interest. The report notes, "[t]he Commission should consider whether rulemaking would be appropriate to prohibit certain conflicts, to require firms to mitigate conflicts through specific action, or to impose specific disclosure and consent requirements." When it comes to duty of care, the study suggests that minimum baseline professional standards should be adopted that could include, for example, specifying what basis a broker-dealer or investment adviser should have in making a recommendation to an investor.

III. <u>Traditional standards of care for Investment Advisers and Broker-Dealers</u>

A. <u>Investment Advisers</u>

Section 202(a)(11) of the Advisers Act defines an "investment adviser" as:

> 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 as part of a regular business, issues or promulgates analyses or reports concerning securities.

Section 202(a)(11)(C) of the Advisers Act excludes from the definition of an investment adviser any broker or dealer that meets the following requirements: (1) the performance of investment advisory services is 'solely incidental' to the conduct of its business as a broker-dealer, and (2) no "special compensation" is received for advisory services.

Investment advisers owe their clients a fiduciary duty of care (SEC v. Capital Gains Research Bureau, Inc., 1963; Transamerica Mortgage Advisors, Inc., 1979). The fiduciary standard that applies to investment advisers encompasses the adviser's entire relationship with its clients and prospective clients (SEC v. Capital Gains Research Bureau, Inc., 1963) and imposes a duty of loyalty and a duty of care.

The duty of loyalty requires a fiduciary to act in the best interests of the client even if doing so may not be in the financial interests of the fiduciary. Under the duty of loyalty, a fiduciary is required to disclose potential conflicts of interest so that the client is aware of those matters where the adviser, either consciously or unconsciously, might render advice which was not in the best interest of the client (SEC v. Capital Gains Research Bureau, Inc., 1963).

The duty of care requires a fiduciary to "make a reasonable investigation to determine that it is not basing its recommendations on materially inaccurate or incomplete information (U.S. Securities and Exchange Commission, 2003). Investment advisers, as fiduciaries, must make suitable and reasonable investment advice to their clients based on the client's financial situation and investment objectives.

B. <u>Broker-Dealers</u>

Traditionally, a broker-dealer has acted as an intermediary between a buyer and seller of securities. Unlike investment advisers, which are subject to a fiduciary standard, broker-dealers have traditionally been subject to a less stringent standard referred to as the "suitability standard." The suitability standard requires broker-dealers to provide suitable investments to customers, but does not require the broker-dealer to act in their best interest (Simon, 2005).

Broker-dealers do, however, have an obligation to deal fairly with customers. Courts have found that broker-dealers make an implicit representation to customers that they will be treated fairly in a manner that is consistent with the standards of the profession (Charles Hughes & Co. v. SEC, 1943). Through various rulemaking initiatives, FINRA (and its predecessor organization, the National Association of Securities Dealers, or NASD) has helped define the duties implied by this fair dealing standard. Among these duties are requirements for broker-dealers to have a reasonable basis for recommendations that are made after considering the customer's financial situation (i.e., a "suitability standard") (NASD Rule 2310); engage in fair and balanced communications with the public (NASD Rule 2210(d)); provide timely and adequate confirmation of transactions; provide account statements (NASD Rule 2340); disclose conflicts of interest (NASD Rule 2720; NASD Rule 3040); receive fair compensation in agency and principal transactions (NASD Rule 2440; FINRA Rule 5110(c)); and give customers an opportunity to resolve disputes through arbitration.

Broker-dealers typically hire agents to provide their services directly to the public. Stockbrokers, for example, are considered agents of a broker-dealer. This agency relationship further complicates matters (and leads to confusion by the public about the varying standards that apply to investment advisers and broker-dealers) because an agent owes his or her primary duty to the principal (which, in this case, would be the broker-dealer). The duty of loyalty owed to the principal (broker-dealer) transcends any duty that the agent may have to a customer while acting in the role of an intermediary. "Even if a non-fiduciary stockbroker wanted to follow the trust standard of law and become a fiduciary to its clients, it cannot do so because of the conflict it has with its broker-dealer. Such contracts require the stockbroker to place the interests of the broker-dealer before the interests of the stockbroker's clients" (Simon, 2005).

While broker-dealers are not subject to the fiduciary standard under federal law, state common law may impose a fiduciary standard on broker-dealers providing services within that state in addition to rules and regulations imposed by the federal government for transactions and services. Courts in four states have chosen to impose an unambiguous fiduciary standard on broker-dealers.

IV. Study Objective

As a response to the regulatory problems and perceived fraud in financial markets that contributed to the financial crisis, Congress passed, and the President signed into law, the Dodd-Frank Act. Prior to the financial crisis, some private self-regulatory organizations, such as Certified Financial Planner Board of Standards, Inc. (CFP Board) sought to distinguish designees from other providers of financial services by holding certificants to a fiduciary standard of care when dealing with clients. These events, along with a perception by lawmakers that higher standards should be applied to providers of financial products and advice, led Congress to call for the completion of a study by the SEC to determine whether it would make sense to impose a unified fiduciary duty of care on both investment advisers and broker-dealers when providing personalized investment advice.

While there has been some recent convergence of the regulatory duties performed by investment advisers and broker-dealers over time, particularly in the area of disclosure, there remain some differences in the scope of services provided by these professionals. Investment advisers have traditionally served higher income/higher net worth clients and are often compensated on an assets under management basis. Depending upon the scope of the engagement, and whether they hold discretion, investment advisers may also hold a duty of care to clients to carefully monitor investment performance. Beginning in the late 1980s and early 1990s, the landscape for the delivery of investment advice began to shift when broker-dealers began to increasingly offer financial advice, relying on the "solely incidental" exemption in the Advisers Act, or becoming dually registered as investment advisers to provide fee-based advisory services. The investment advice provided on the brokerage side, however, tends to be episodic and focused on specific products and transactions that are suitable for a given client. Broker-dealer agents are usually compensated on a commission basis, and traditionally do not owe customers an ongoing duty to monitor their client's financial position. Broker-dealers have claimed to provide lower-cost advisory services, offset by transaction fees, for customers who do not wish to pay, or cannot afford to pay, the higher direct fees charged by investment advisers.

Due, in part, to the imposition of the suitability (as opposed to fiduciary) standard on broker-dealers, the current debate over the costs of providing advisory services to retail customers has focused on the potential economic effects of broker-dealers being held to the higher fiduciary standard of care. The brokerage industry argues that the imposition of a fiduciary standard will result in an increased risk of a fiduciary breach that would have the effect of increasing the compliance and liability costs of providing traditional broker-dealer services, and, consequently, may make those services too expensive for many lower or middle income clients (Headley, 2011).

Further, while imposing a fiduciary standard of care may provide additional protections for brokerage customers, critics assert that the imposition of such a standard may result in some customers losing access to financial advice if the cost of that advice rises due to the imposition of the standard, or, alternatively, some customers may find that they will have to pay more for the investment advice they receive without experiencing a significant change in service due to the increased regulatory and liability costs imposed by regulation.

In order to test claims that the brokerage industry and their customers would be adversely affected by the imposition of a stricter fiduciary standard, this study surveyed registered representatives (brokers) of broker-dealers in states that impose a fiduciary duty on the provision of investment advice to retail investors, and in states that did not impose such a duty. The survey avoided brokers who are dually registered as investment adviser agents and who, in that capacity, provide fiduciary investment advice. If presence of a fiduciary duty for brokers results in higher costs associated with that standard, it would suggest that states that impose the higher fiduciary standard have a lower saturation of brokers to households within that state. This would imply that there is an additional service cost attached to imposition of the fiduciary standard by reducing the number of service providers for lower or middle-income customers.

V. Differentiating State Law

States were divided into three categories: 1. states that unambiguously apply a fiduciary standard to brokers in that state; 2. states that unambiguously apply no fiduciary standards to brokers; and 3. states where there is evidence of a limited fiduciary standard applied to brokers.

Four states have imposed an unambiguous fiduciary standard on broker-dealers (fiduciary states). These states are California, Missouri, South Dakota, and South Carolina. California, Missouri, and South Dakota courts expressly impose a fiduciary duty on broker-dealers.

California courts, for example, have held that a broker's fiduciary duty requires that he or she act in the highest good faith toward the customer (Hobbs v. Bateman Eichler, Hill Richards, Inc., 1985). Missouri courts have held that, "stockbrokers owe customers a fiduciary duty. This fiduciary duty includes at least these obligations: to manage the account as directed by the customer's needs and objectives, to inform of risks in particular investments, to refrain from self-dealing, to follow order instructions, to disclose any self-interest, to stay abreast of market changes, and to explain strategies" (State ex rel Paine Webber v. Voorhees, 1995). South Dakota courts have held that securities brokers owe the same fiduciary duties to customers as those owed by real estate brokers, including a duty of utmost good faith, integrity, and loyalty, and a duty to act primarily for the benefit of another (Dismore v. Piper Jaffray, Inc., 1999). While South Carolina courts have not expressly stated that broker-dealers must live up to a fiduciary standard, the courts have imposed duties commensurate with those required when a fiduciary duty applies, including a duty to refrain from acting contrary to a customer's best interest, avoiding fraud, and communicating information to the customer that would be in the customer's advantage (Cowburn v. Leventis, 2005). South Carolina courts have clearly imposed a duty of care commensurate with the duty required by a fiduciary that exceeds the suitability standard that applies under federal law to broker-dealers.

States that do not impose a fiduciary standard on broker-dealers are Arizona, Arkansas, Colorado, Hawaii, Massachusetts, Minnesota, Mississippi, Montana, New York, North Carolina, North Dakota, Oregon, Washington, and Wisconsin.

Courts in Arkansas, Hawaii, Massachusetts, Montana, and Washington have expressly stated that, under state law, a fiduciary duty does not exist between a client and a broker-dealer. The U.S. Federal District Court, and the U.S. Court of Appeals for the 8th Circuit have held that under

Arkansas law, no fiduciary duty is owed by a commodities broker to a nondiscretionary account holder (Greenwood v. Dittmer, 1985). Likewise, the Federal District Court of Hawaii has concluded that Hawaii law does not impose a fiduciary duty on brokers (Unity House, Inc. v. North Pacific Inv., Inc., 1996). Courts interpreting Montana and Washington law have expressly stated that a broker–dealer does not owe a fiduciary duty to a non-discretionary account holder (Willems v. U.S. Bancorp Piper Jaffray, Inc., 2005; Chor v. Piper, Jaffray & Hopwood, 1993; Sherry v. Dierks, 1981). Massachusetts courts have expressly stated that "Under Massachusetts law, a 'simple' broker-customer relationship is not fiduciary in nature…" (Pastos v. First Albany, 2001; Vogelaar v. H.L. Robbins & Co., 1965).

Courts in Arizona, Colorado, Mississippi, New York, North Carolina, North Dakota, and Oregon have all concluded that broker-dealers do not owe a fiduciary duty to holders of non-discretionary accounts (SEC v. Raucher Pierce Refsnes, Inc., 1998; Rhoads v. Harvey Publications, Inc., 1984; Hudson v. Wilhelm, 1987; Puckett v. Rufenacht, Bromagen & Hertz, 1991; Fesseha v. TD Waterhouse Investor Servs., 2003; Sterner v. Penn, 2003; Ray E. Friedman & Co. v. Jenkins, 1984; Berki v. Reynolds Securities, Inc., 1977; Wallace v. Hinkle Northwest, 1986). In Minnesota and Wisconsin, state law provides that a broker does not owe a fiduciary duty to customers absent a special agreement between the parties (MERF v. Allison-Williams Co., 1993; Rude v. Larson, 1973; Merrill Lynch v. Boeck, 1985).

The remaining states (Alabama, Alaska, Connecticut, Delaware, Florida, Georgia, Idaho, Illinois, Indiana, Iowa, Kansas, Kentucky, Louisiana, Maine, Maryland, Michigan, Nebraska, New Hampshire, New Jersey, New Mexico, Nevada, Ohio, Oklahoma, Pennsylvania, Rhode Island, Tennessee, Texas, Utah, Vermont, Virginia, West Virginia, and Wyoming) impose either a limited fiduciary standard, or the courts have interpreted state law to impose duties that appear to be fiduciary in nature. In this study, these states are referred to as quasi-fiduciary states. Quasi-fiduciary states impose standards that exceed the suitability standard set forth under FINRA rules, but do not expressly classify broker-dealers as fiduciaries. The duties imposed, and the manner in which they are imposed, vary among these states. In Alaska, for example, courts have found that fiduciary duties arise "when one imposes a special confidence in another, so that the latter, in equity and good conscience, is bound to act in good faith and with due regard to the interests of the one imposing the confidence" (Enders v. Parker, 2003). While the Enders court did not specifically consider whether a fiduciary duty is imposed on a broker-dealer, the court's standard for imposing a fiduciary duty could reasonably be interpreted to create a duty for a broker-dealer in some circumstances. Other states, such as Connecticut, refrain from imposing an express fiduciary duty, but did find an agency relationship between a broker and a client which required the broker to exercise "reasonable skill, care, and diligence" (Precision Mechanical v. T.J.PFund, 2003). Connecticut's approach is intriguing in that an agency relationship exists with both the registered representative's employer (the broker-dealer) and with the customer. Connecticut law, as currently expressed, cannot impose a fiduciary duty on registered representatives due to the inherent conflict of interest created by the state's imposition of a customer-representative agency relationship which suggests that the registered representative serves two masters, not one. Iowa courts have not traditionally imposed a fiduciary duty on a broker-client relationship, but do so when certain circumstances exist, such as when the client lacks prior investment experience, the advice offered by the broker-dealer is significant, the client relies (to his detriment) on the advice provided by the broker dealer, and the broker-dealer was aware that the client had not read any literature concerning the subject (McCracken v. Edward D. Jones & Co., 1989).

States that impose a limited fiduciary duty include Delaware, Florida, Georgia, Illinois, Kansas, Louisiana, Maryland, Michigan, Ohio, Pennsylvania, Tennessee, and Texas. Almost all of these states impose a standard higher than the suitability standard imposed by FINRA for non-discretionary accounts. Louisiana does not expressly impose a standard of conduct higher than the suitability standard, but does require a court to consider a variety of circumstances when determining whether a higher standard should exist. The items that Louisiana courts must consider include the relationship between the broker-dealer and client, the nature of the account, and the sophistication of the customer (Beckstrom v. Parnell, 1998).

VI. Criticisms of the impact of imposing a fiduciary standard

Under current law, investment advisers are subject to a fiduciary standard under the Advisers Act, while broker-dealers are subject to a suitability standard. Differing client characteristics have resulted in different business models used by investment advisers and broker-dealers to deliver cost effective advice to their clients. Imposing a uniform fiduciary standard on both investment advisers and broker-dealers may have unintended consequences.

Some in the brokerage industry have argued that the imposition of fiduciary regulation will lead to reduced consumer access to financial advice, particularly among middle-class households that may not have access to investment advisers. Many broker-dealers provide financial services other than the sale of securities to their clients, including insurance products and brokerage services to qualified retirement plans. The president of the National Association of Insurance and Financial Advisors (NAIFA) testified before the House Committee on Financial Services that broker-dealers are typically subject to both additional state and federal regulation for these services, and these regulations generally provide constraints on behaviors that may be considered abusive (Headley, 2011).

Imposing the higher fiduciary standard than currently applies to investment advisers may increase the compliance costs of broker-dealers. A study conducted by NAIFA in 2010 found that an unintended consequence of imposing a uniform fiduciary standard would be to "negatively impact product access, product choice, and affordability of customer services for those customers who are in most need of these services" (Headley, 2011). Specifically, the study indicated that imposition of a uniform fiduciary standard may "create the potential for market disruption and reduced choices for investors when it comes to who they work with and how they pay for services" (National Association of Insurance and Financial Advisors (in Partnership with LIMRA), 2010). The NAIFA study indicated that most of its members are "concerned that the additional regulatory requirements and potential legal implications of a fiduciary standard could significantly increase their compliance costs" (Headley, 2011; National Association of Insurance and Financial Advisors (in Partnership with LIMRA), 2010). In the NAIFA study, sixty-five (65) percent of NAIFA members indicated that if compliance costs rose by 15 percent, they would limit their practice to affluent clients only (31 percent of those surveyed), would not offer securities to their clients (20 percent of those surveyed), or would increase fees for their clients (14 percent of those surveyed) (Headley, 2011).

An SEC staff study indicated that investors "generally were satisfied with their financial professionals" (Staff of the U.S. Securities and Exchange Commission, 2011), but that customers are confused with the varying standards that apply to different types of financial advisers and based on this conclusion recommended the adoption of a uniform fiduciary standard (Staff of the U.S. Securities and Exchange Commission, 2011). While the industry raised concerns that imposing a uniform standard that increases compliance costs for broker-dealers may result in limited access to suitable investment advice for middle-income clients, the SEC staff noted the possibility that the change in standards might result in reduced administrative and compliance costs (Staff of the U.S. Securities and Exchange Commission, 2011).

Opponents of the fiduciary standard are often criticized for having no data to substantiate claims about increased costs that may arise upon imposition of a uniform fiduciary standard (Consumer Federation of America, 2011). In particular, proponents of a uniform fiduciary standard assert that "claims about increased liability costs associated with a fiduciary duty are...unsupported and ignore the legal environment in which brokers currently operate" (Consumer Federation of America, 2011) because "the SEC proposal makes clear that it intends to provide extensive guidance to assist brokers in implementing the fiduciary standard" (Consumer Federation of America, 2011). Proponents of a uniform standard claim that the SEC proposal "would not require brokers to charge fees", and that the proposal preserves "the ability of brokers to offer transaction-based advice...[while] at the same time...rais[ing] the standard that applies to those transaction based recommendations" (Consumer Federation of America, 2011).

Imposing a fiduciary standard on transaction-based advice may increase the potential for legal liability of the registered representative, requiring the broker to be compensated for that additional risk. NAIFA members have expressed concern that the increased duties they owe transactional clients under a fiduciary standard may result in potential legal implications that increase their cost of doing business (National Association of Insurance and Financial Advisors (in Partnership with LIMRA), 2010).

VII. Methods

In order to estimate how the imposition of a stricter universal fiduciary standard will impact the provision of financial advice within the brokerage industry, we obtained the names and addresses of 544,000 registered representatives active in November 2011 and sorted them into categories based on the application of a fiduciary standard. There are four states that apply a strict fiduciary standard, 14 that apply a limited fiduciary standard, and 32 states (and the District of Columbia) that apply no fiduciary standard.

Our objectives were to assess perceived differences in business conduct among registered representatives sorted by fiduciary regulation and to assess the market saturation (representatives as a proportion of total households) of registered representatives among these states. To assess whether registered representatives' business conduct differs in states that apply a strict fiduciary standard, we developed a survey among a sample of registered representatives in states that apply no fiduciary standard and states that provide a strict fiduciary standard. The survey was conducted in the months of November and December, 2011. Participants were drawn randomly from both categories of states and were asked twelve questions. These questions were based on brokerage industry statements and testimony before Congress suggesting that a stricter fiduciary standard will result in differences in ability to serve moderate wealth customers, to offer a variety of products, to provide product recommendations that are in the best interest of their customers, and whether representatives experience a greater compliance burden. Representatives were phoned in their offices and those dually registered as investment advisers are excluded from the analysis since we are unable to differentiate whether their responses relate to their activities conducted under a fiduciary or suitability regime.

Broker-dealers in fiduciary and non-fiduciary states were asked the following questions:

1. Are you a registered investment adviser? (If so, survey is over.)
2. What percentage of your clients have incomes of less than $75,000?

3. What percentage has investable assets of over $750,000?
4. Are you able to serve the financial needs of low to moderate wealth clients?
5. Do your state's security regulations limit your ability to recommend a broad range of financial products?
6. Do you offer your clients a choice of financial products that meet their financial needs and objectives?
7. Do you provide advice tailored to the specific needs of your clients?
8. Do you feel that less affluent clients avoid obtaining your services due to cost?
9. Are you able to recommend products that provide a commission?
10. How significant is the cost of compliance?
11. Do you feel that you make product recommendations that are in the best interest of your client?
12. Among the following options, which do you consider to be the most important single factor in pricing your investment advice to clients: competition in the marketplace, firm brand, personal qualifications, legal and compliance burden, or other?

In order to provide insight into whether the imposition of stricter fiduciary standards leads to reduced supply, we compared the saturation of registered representatives within the total population of states sorted into the three fiduciary categories (strict, limited and no fiduciary standard). Individuals complete examinations conducted by FINRA in order to become registered representatives that are able to facilitate transaction with individual investors. Completion of the Series 6 and Series 7 examinations is necessary to sell, respectively, investment company products and individual securities, to the public. Only registered representatives who have completed Series 6 or Series 7 examinations were included in the analysis.[1] We provide both a descriptive comparison of saturation among states and a multivariate analysis that includes dummy variables for strict fiduciary and non-fiduciary standards with limited fiduciary as the reference category. Due to the small sample size (50 states and the District of Columbia), we include one control variable to account for the log of mean household income within the state.

[1] This constraint excludes less than 5% of the original sample and has no impact on the empirical results.

New York housed five of the 17 largest broker-dealer firms in the United States in 2011 (Investment News, 2012). The saturation of brokers within New York state is more than three times the national average and twice as high as the second largest state (Colorado). Since New York is the traditional center of the brokerage industry and may include a large number of registered representatives not primarily engaged in selling securities directly to individual clients, we include descriptive statistics with and without New York state and include an additional multivariate analysis with a dummy variable to control for the New York effect.

VIII. Results

Descriptive statistics summarizing the responses received from a random survey of 207 registered representatives in the four strict fiduciary states and the 14 non-fiduciary states are presented in Table 1. The percentage of clients who have an income of less than $75,000 is statistically equal between both groups, and there is no statistically significant difference in either the percentage of high wealth clients or in the percentage of brokers who believe they serve the needs of low and moderate wealth clients. Nearly all respondents believe they are able to provide products and advice that meet the needs of customers. The percent who respond that they are able to recommend commission products is 88.5% in strict fiduciary states and 88.2% in non-fiduciary states. The largest percentage point difference among any of the questions is whether the cost of compliance is significant. 70.9% of respondents in fiduciary states felt the costs were significant compared to 61.9% in non-fiduciary states. This difference, and that of all other questions in the survey, was not statistically significant.

Mean rates of broker saturation calculated as the number of registered representatives divided by the number of households within the state are presented in Table 2. There is a wide range in saturation rates among states from a low of 1.31 per 1,000 households in New Mexico to a high of 13.41 in New York. Average saturation rates are lowest among states with a limited fiduciary standard (3.81) and highest among states with no fiduciary standard (6.33). However, the saturation rates were nearly identical among fiduciary categories when New York is excluded from the non-fiduciary states. Saturation rates are 3.96 for strict fiduciary states, 3.81 for limited fiduciary, and 4.04 for non-fiduciary states.

We then take Missouri, an average-sized state with a fiduciary standard, and compare it with other states that have a population between 2 and 3 million households (Table 3). The broker saturation rate in Missouri (2.65) is equal to that of Tennessee (a limited fiduciary state) and comparable to non-fiduciary states with similar income levels (Arizona is 3.12, Washington is 2.54). Other states with higher incomes have higher saturation rates.

In order to control for state saturation differences that may be caused by differences in income within states, we run a regression modeling individual state saturation rate as a function of fiduciary status and log household income. Results (Table 4) show that there is no statistical difference in saturation rates among fiduciary and non-fiduciary states relative to the reference group of limited fiduciary states. When a dummy variable is included to account for the elevated saturation within New York state, the coefficient suggests that the saturation rate in New York is 8.3 points higher than the predicted rate. Fiduciary status variables remain statistically insignificant.

IX. Conclusions

This study explores the regulation of registered representatives of broker-dealers in order to estimate whether the proposed application of a universal fiduciary standard will have a significant impact on the financial adviser industry. We take advantage of differences in the application of a fiduciary standard to representatives among states in order to test whether representatives already subject to a stricter fiduciary requirement are affected by the higher standard. We conduct a survey of 207 representatives within the four states that apply a strict fiduciary standard and the 14 states that apply no fiduciary standard and find no statistical differences between the two groups in the percentage of lower-income and high-wealth clients, the ability to provide a broad range of products including those that provide commission compensation, the ability to provide tailored advice, and the cost of compliance.

We then compare the ratio of registered representatives to total households among states within the three fiduciary regimes. When New York (which houses a disproportionate proportion of broker-dealer firms) is excluded from the non-fiduciary states, the saturation rate is almost identical between fiduciary, limited fiduciary and non-fiduciary states. A comparison of a moderate size state with strict fiduciary regulation (Missouri) with non-fiduciary and limited-fiduciary states of a similar population suggests a strong similarity among states with similar incomes.

A multivariate analysis of broker saturation that controls for fiduciary and non-fiduciary regulation as well as state mean income yields no significant fiduciary effect even with New York included as a non-fiduciary state. The addition of a dummy variable to account for the New York effect suggests that New York's saturation rate is inflated by 8.3 representatives per thousand households.

Empirical results provide no evidence that the broker-dealer industry is affected significantly by the imposition of a stricter legal fiduciary standard on the conduct of registered representatives. The opposition of the industry to the application of stricter regulation suggests that agency costs that exist when brokers are regulated according to suitability are significant.

Imposition of a universal fiduciary standard among financial advisers may result in a net welfare gain to society, and in particular to consumers who are ill equipped to reduce agency costs on their own by more closely monitoring an adviser with superior information, although this will likely occur at the expense of the broker-dealer industry. These results provide evidence that the industry is likely to operate after the imposition of fiduciary regulation in much the same way it did prior to the proposed change in market conduct standards that currently exist for brokers.

Works Cited

Beckstrom v. Parnell, 730 So2d 932 (La. App. Ct. 1998) (Louisiana Appellate Court 1998).

Berki v. Reynolds Securities, Inc., C.A. 415-354, CCH Par 71,365 (Or. Feb. 10, 1977) (1977).

Charles Hughes & Co. v. SEC, 139 F.2d 434 (2d Cir. 1943), cert denied, 321 U.S. 786 (1944) (U.S. Court of Appeals for the Second Circuit 1943).

Chor v. Piper, Jaffray & Hopwood, 261 Mont. 143, 862 P.2d 26 (1993) (1993).

Consumer Federation of America. (2011). *Response to Arguments about Fiduciary Duty "Unintended Consequences"*. Consumer Federation of America (CFA).

Cowburn v. Leventis, CCH Par 75,542 (S. Ca. Ct. App. May 16, 2005) (South Carolina Court of Appeals May 16, 2005).

Cummings, B. & Finke, M., The Economics of Fiduciary Investment Advice, Available at SSRN: http://ssrn.com/abstract=1701181 (2010).

Dismore v. Piper Jaffray, Inc., 593 N.W.2d 41 (S.D. 1999) (1999).

Enders v. Parker, 66 P.3d 11 (Ak. 2003) (2003).

Fesseha v. TD Waterhouse Investor Servs., 305 A.D.2d 268 (N.Y. App. Ct. 2003) (New York Appellate Court 2003).

FINRA Rule 5110(c). (n.d.).

Frankel, T., Fiduciary Law, 71, California Law Review, 795, 800 (1983).

Greenwood v. Dittmer, 776 F.2d 785 (8th Cir. 1985) (U.S. Court of Appeals for the 8th Circuit 1985).

Headley, T. (2011, September 13). Ensuring Appropriate Regulatory Oversight of Broker-Dealers and Legislative Proposals to Improve Investment Adviser Oversight. *Testimony offered to the House Committee on Financial Services, Subcommittee on Capital Markets and Government Sponsored Enterprises* . Washington, DC.

Hobbs v. Bateman Eichler, Hill Richards, Inc., 165 Cal. App. 3d 174, 210 Cal. Rptr. 387 (Cal. Ct. App. 1985) (California Court of Appeals 1985).

Hudson v. Wilhelm, 651 F.Supp. 1062 (D. Col. Jan 12, 1987) (Federal District Court of Colorado January 12, 1987).

Hung, A. A., Clancy, N., Dominitz, J., Talley, E., Berrebi, C., & Suvankulov, F. (2008). *Investor and industry perspectives on investment advisers and broker-dealers.* RAND Corporation.

Investment News, Broker-dealer rankings, available at http://www.investmentnews.com/section/broker-dealer-data-rankings

Jensen, M., & Meckling, W. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics , 3* (4), 305-360.

McCracken v. Edward D. Jones & Co., 445 N.W.2d 375 (Iowa App. Ct. 1989) (Iowa Appellate Court 1989).

MERF v. Allison-Williams Co., 508 N.W.2d 805 (Minnesota Court of Appeals 1993).

Merrill Lynch v. Boeck, 127 Wisc.2d 127 (1985) (1985).

NASD Rule 2210(d). (n.d.).

NASD Rule 2310. (n.d.).

NASD Rule 2340. (n.d.).

NASD Rule 2440. (n.d.).

NASD Rule 2720. (n.d.).

NASD Rule 3040. (n.d.).

National Association of Insurance and Financial Advisors (in Partnership with LIMRA). (2010). *Facts at a Glance: Results from a LIMRA survey of NAIFA members.* Falls Church, Virginia: NAIFA.

Pastos v. First Albany, 433 Mass. 323, 741 F.2d 841 (2001) (2001).

Precision Mechanical v. T.J.PFund, CA 90-0416692, 2003 Ct. Sup. 14518 (Conn. Sup. Ct. Dec. 22, 2003) (Connecticut Superior Court December 22, 2003).

Puckett v. Rufenacht, Bromagen & Hertz, 587 So2d 273 (Miss. 1991) (1991).

Ray E. Friedman & Co. v. Jenkins, 738 F.2d 251 (8th Cir. 1984) (U.S. Court of Appeals for the 8th Circuit 1984).

Rhoads v. Harvey Publications, Inc., 700 P.2d 840 (Ariz. Ct. App. 1984) (Arizona Court of Appeals 1984).

Rude v. Larson, 296 Minn. 518, 207 N.W.2d 709 (1973) (Minnesota Court of Appeals 1973).

SEC v. Capital Gains Research Bureau, Inc., 375 U.S. 180 (U.S. Supreme Court 1963).

SEC v. Raucher Pierce Refsnes, Inc., 98 Fed.Sec.L.Rep. (CCH) Par 90,284 (D Ariz. Aug 10, 1998) (Federal District Court of Arizona 1998).

Sherry v. Dierks, 29 Wn. App. 433, 628 P.2d 1336 (1981) (1981).

Simon, W. S. (2005). *The Critical Difference Between a Stockbroker and Registered Investment Advisor.*

Staff of the U.S. Securities and Exchange Commission. (2011). *Study on Investment Advisers and Broker-Dealers.* Washington, DC.

State ex rel Paine Webber v. Voorhees, 891 S.W.2d 126 (Mo. 1995) en banc (1995).

Sterner v. Penn, 159 N.C. App. 626 (2003) (North Carolina Appellate Court 2003).

Transamerica Mortgage Advisors, Inc., 444 U.S. 11 (U.S. Supreme Court 1979).

U.S. Securities and Exchange Commission. (2003, Jan 31). Proxy Voting by Investment Advisers, Investment Advisers Act Release No. 2106.

Unity House, Inc. v. North Pacific Inv., Inc., 918 F.Supp. 1384 (D. Ha. 1996) (Federal District Court of Hawaii 1996).

Vogelaar v. H.L. Robbins & Co., 348 Mass. 787, 204 N.E.2d 461 (1965) (1965).

Wallace v. Hinkle Northwest, 79 Or.App. 177 (Or. App. 1986) (Oregon Appellate Court 1986).

Willems v. U.S. Bancorp Piper Jaffray, Inc., 326 Mont. 103, 107 P.3d 465 (2005) (2005).

Table 1: Mean and frequency comparison of registered representatives

Question	Fiduciary States	Non-Fiduciary States	Difference (Fiduciary – NF)	P-Value Equal	DF
% clients income < $75,000	28.0%	27.9%	0.1%	0.982	174
% clients inv assets > $750,000	29.5%	34.5%	-5.0%	0.261	183
Serve needs of low/mod wealth	78.9%	79.8%	-0.9%	0.878	202
Regulation limits product range	21.3%	17.4%	3.9%	0.486	198
Products meet client needs	95.8%	97.3%	-1.5%	0.561	207
Advice tailored to client needs	91.7%	90.1%	1.6%	0.695	207
Less affluent avoid due to cost	23.6%	29.2%	-5.6%	0.374	195
Able to recommend commission	88.5%	88.2%	0.3%	0.936	206
Cost of compliance significant	70.9%	61.9%	9.0%	0.190	191
Act in best interest of client	97.8%	96.3%	1.5%	0.526	202

Table 2: Broker Saturation Rates by States

	Registered Representatives	Households (000s)	Saturation
Fiduciary States			
California	56,945	12,392	4.60
Missouri	6,244	2,355	2.65
South Carolina	2,667	1,753	1.52
South Dakota	737	317	2.32
Total Fiduciary	**69,120**	**16,817**	**3.96**
Non-Fiduciary States			
New York	96,862	7,221	13.41
North Carolina	15,094	3,666	4.12
Washington	6,605	2,601	2.54
Massachusetts	16,207	2,521	6.43
Arizona	7,280	2,333	3.12
Wisconsin	10,164	2,282	4.45
Minnesota	8,644	2,093	4.13
Colorado	14,168	1,942	7.30
Oregon	5,291	1,506	3.51
Arkansas	1,787	1,120	1.60
Mississippi	1,728	1,085	1.59
Hawaii	974	443	2.19
Montana	949	404	2.35
North Dakota	1,049	278	3.77
Total Non-Fiduciary	**186,802**	**29,501**	**6.33**
Total W/O New York	**89,940**	**22,279**	**4.04**
Other States			
Texas	39,005	8,666	4.50
Florida	33,968	7,087	4.79
Pennsylvania	24,223	4,952	4.89
Illinois	17,258	4,768	3.62
Ohio	12,385	4,544	2.73
Michigan	8,130	3,815	2.13
Georgia	7,973	3,488	2.29
New Jersey	24,146	3,176	7.60
Virginia	7,836	2,986	2.62
Indiana	8,339	2,471	3.37
Tennessee	6,539	2,454	2.66
Maryland	9,781	2,122	4.61
Alabama	2,701	1,823	1.48
Kentucky	5,404	1,684	3.21
Louisiana	4,789	1,678	2.85
Oklahoma	3,837	1,429	2.68
Connecticut	12,682	1,361	9.32

Iowa	3,190	1,219	2.62
Kansas	2,691	1,106	2.43
Nevada	1,723	984	1.75
Utah	5,611	873	6.42
New Mexico	996	759	1.31
West Virginia	1,275	742	1.72
Nebraska	2,583	715	3.61
Idaho	1,727	574	3.00
Maine	1,291	550	2.35
New Hampshire	2,818	515	5.47
Rhode Island	2,074	408	5.08
Delaware	1,402	331	4.23
District of Columbia	1,872	256	7.31
Vermont	836	256	3.27
Alaska	593	251	2.36
Wyoming	568	219	2.58
Total Other States	**260,246**	**68,278**	**3.81**

Table 3: Comparison of Broker Saturation

This table compares characteristics of Missouri, a state that regulates brokers as fiduciaries, with all other states that have between 2 and 3 million households.

	State Regulation	Reps/ Hhlds	Median Income	Mean Income	% High Income	% College Education
Missouri	Fiduciary	2.65	45,829	60,760	5.36	25.31
Washington	Non-Fid.	2.54	56,911	73,854	8.99	31.02
Massachusetts	Non-Fid.	6.43	63,961	85,865	13.52	38.54
Arizona	Non-Fid.	3.12	49,214	65,552	6.68	26.12
Wisconsin	Non-Fid.	4.45	50,814	64,463	5.55	25.88
Minnesota	Non-Fid.	4.13	56,456	72,850	8.35	31.59
Virginia	Other	2.62	61,090	82,369	12.83	33.92
Indiana	Other	3.37	46,529	60,275	4.90	22.70
Tennessee	Other	2.66	42,612	58,360	5.37	22.92
Maryland	Other	4.61	70,017	90,800	15.18	35.58

Table 4: Broker Saturation Regression Analysis

Panel A
Dependent variable is the ratio of registered representatives to households within 50 states and the District of Columbia. Log income is the natural log of mean household income for each state. Fiduciary is a dummy variable indicating the four states that hold representatives to a fiduciary standard, and non-fiduciary includes the 14 states that do not apply a fiduciary standard to representatives. The omitted reference category is the remaining 33 states (and DC) that do not unambiguously treat representatives as either fiduciaries or non-fiduciaries.

Variable	Coefficient	P-Value
Fiduciary	-0.488	0.601
Non-Fiduciary	0.759	0.180
Log Income	8.941	0.000
Adjusted R-Square	0.39	

Panel B
Adds a dummy variable indicating New York State

Variable	Coefficient	P-Value
Fiduciary	-0.542	0.447
Non-Fiduciary	-0.154	0.726
Log Income	7.741	0.000
New York Dummy	8.290	0.000
Adjusted R-Square	0.65	

41

COMMENT TO THE DEPARTMENT OF LABOR ON A PROPOSED RULE REGARDING FIDUCIARY STATUS UNDER ERISA

Daniel R. Fischel and Todd D. Kendall

April 12, 2011

COMPASS LEXECON

I. INTRODUCTION AND SUMMARY

1. I, Daniel R. Fischel, am President of Compass Lexecon, a consulting firm that specializes in the application of economics to a variety of legal and regulatory issues. I am also Professor of Law and Business at Northwestern University School of Law and Kellogg School of Management, as well as the Lee and Brena Freeman Professor of Law and Business Emeritus at The University of Chicago Law School. I served previously as Dean of The University of Chicago Law School, as Director of the Law and Economics Program at The University of Chicago Law School, and as Professor of Law and Business at The University of Chicago Graduate School of Business.

2. In the past, I have served as a consultant or adviser on economic issues to, among others, the United States Department of Labor, the United States Securities and Exchange Commission, the United States Department of Justice, the National Association of Securities Dealers, the New York Stock Exchange, the Chicago Board of Trade, the Chicago Mercantile Exchange, the New York Mercantile Exchange, the Federal Deposit Insurance Corporation, the Resolution Trust Corporation, and the Federal Trade Commission.

3. Much of my research and teaching have addressed the law and economics of financial markets, including the proper role of pension plan fiduciaries.[1] I have published approximately fifty articles in leading legal and economics journals and am coauthor, with Judge Frank Easterbrook of the Seventh Circuit Court of Appeals, of the book *The Economic Structure of Corporate Law* (Harvard University Press). Courts of all levels, including the Supreme Court of the United States, have cited my articles as authoritative.

4. I am a member of the American Economic Association and the American Finance Association. I am also a former member of the Board of Directors of the Center for the Study of the

1. See, e.g., Daniel Fischel and John H. Langbein (1988) "ERISA's Fundamental Contradiction: The Exclusive Benefit Rule", *U. Chi. L. Rev.* 55:1105-60, and Frank H. Easterbrook and Daniel R. Fischel (1993) "Contract and Fiduciary Duty", *Journal of Law & Economics* 36:425-46.

Economy and the State at The University of Chicago, and former Chairman of the American Association of Law Schools' Section on Law and Economics.

5. I, Todd D. Kendall, am a Senior Economist at the aforementioned firm, Compass Lexecon. Previously, I served on the faculty of the economics department at Clemson University. I have published approximately a dozen articles in academic economics journals and collected volumes on the topic of applied economic theory, and which employ statistical and econometric methods. I have been employed at Compass Lexecon since 2008, during which time I have consulted on a wide range of regulatory, litigation, merger and other business matters involving brokerage services, banks, securities exchanges, and other industries. I am a member of the American Economic Association and the Econometric Society.

6. We understand that the Department of Labor (the "DOL") is currently considering a rule (the "Proposed Rule") that would broaden the circumstances under which a person is considered to be a "fiduciary" under the Employee Retirement Income Security Act and the Internal Revenue Code.

7. For the purposes of our analysis, we have been asked by counsel for Primerica to consider the consequences if the Proposed Rule led to fiduciary status where none currently exists for certain companies and their representatives ("commission-based brokers") who provide brokerage and other services to investors regarding individual retirement accounts ("IRAs"), and who receive certain types of compensation paid by third parties providing financial products in connection with IRAs.[2] We have also been asked to assume that, if these commission-based brokers were deemed to be fiduciaries, they would significantly limit their receipt of this compensation.

8. We have been asked by counsel for Primerica to (a) identify any significant costs or benefits of the Proposed Rule other than those presented in the cost-benefit analysis presented by the

2. Throughout this report, we focus on financial services provided for IRA investors, although we understand that the Proposed Rule may also impact service providers with respect to other investment products, such as Coverdell education savings accounts.

DOL in this matter[3]; and (b) evaluate whether the evidence provided by the DOL, or other available evidence, is sufficient to conclude that the benefits of the Proposed Rule outweigh the costs.

9. Our main conclusions are as follows:

- There are several important costs not quantified in the DOL cost-benefit analysis that would likely result from the Proposed Rule, leading to significantly higher costs than estimated by the DOL.
- The evidence presented by the DOL supporting alleged benefits from the Proposed Rule does not provide a sufficient basis to conclude that these benefits would be large enough to outweigh the costs.
- A review of economic theory and available evidence regarding the IRA investment services industry does not support a conclusion that the Proposed Rule would generate benefits large enough to outweigh the costs.

We explain the basis on which we came to these conclusions in the following three sections of this Comment.

10. We understand that Oliver Wyman has also performed a separate analysis of the Proposed Rule, based on proprietary data from 12 IRA brokerage firms, and came to similar conclusions.[4]

II. THERE ARE SEVERAL IMPORTANT COSTS NOT QUANTIFIED IN THE DOL COST-BENEFIT ANALYSIS THAT WOULD LIKELY RESULT FROM THE PROPOSED RULE, LEADING TO SIGNIFICANTLY HIGHER COSTS THAN ESTIMATED BY THE DOL.

11. The DOL cost-benefit analysis estimates the monetized costs of the Proposed Rule at between \$15.6 million and \$16.7 million.[5] This figure is based on an estimate of the legal costs that financial service providers would incur for a compliance review of their books of business under the Proposed Rule. Undoubtedly, financial service providers and their representatives would incur significant compliance costs in complying with new regulations. We have not attempted to fully

---

3. 75 Fed. Reg. 62570-8 (2010-10-22) (hereafter, "DOL cost-benefit analysis").
4. Oliver Wyman (2011) "Assessment of the Impact of the Department of Labor's Proposed 'Fiduciary' Definition Rule on IRA Consumers", April 12, 2011.
5. DOL cost-benefit analysis, *supra*, at 65274, Table 2.

evaluate the accuracy of the monetized cost estimate provided by the DOL with respect to legal compliance costs; however, we believe it may be understated for at least two reasons. First, this estimate does not appear to incorporate the potentially very large additional legal costs financial services firms would likely incur to defend against litigation associated with their new status as fiduciaries, even after a full review of their books of business, or to purchase fiduciary liability insurance. Second, the DOL estimate is based on an assumption that affected firms would require a certain number of hours of legal professional time, valued at $119 per hour.[6] We understand that this rate is substantially lower than the rate assumed by other government agencies. For instance, using a rate consistent with that assumed by the Securities and Exchange Commission ("SEC"), which recently valued legal professional time at $354 per hour[7] would increase the DOL's estimate of the monetized costs of the Proposed Rule to between $46.5 million and $49.7 million, using the same discount rates employed by DOL.

12. In any case, the DOL cost estimate does not incorporate several important costs besides legal services that would likely result as a consequence of the Proposed Rule. Specifically, the DOL cost-benefit analysis does not quantify likely potential costs of the Proposed Rule due to (a) higher certification requirements for IRA service providers, (b) increased expenses paid by IRA investors, and (c) lower returns on investors' retirement funds. As we discuss below, the potential size of these effects is large. In addition, there may be other costs besides those that we describe here; however, we believe these three illustrate the potential for costs associated with the Proposed Rule significantly higher than estimated by the DOL.

---

6. DOL cost-benefit analysis, *supra*, at 65274.
7. 76 Fed. Reg. 15003.

A. Costs Due to Higher Certification Requirements for IRA Service Providers Would Likely Rise Significantly Due to the Proposed Rule.

13. We understand that many representatives of broker-dealer firms that currently provide services to IRA investors do not currently hold the certifications necessary to operate as fee-based investment advisers, and that if the Proposed Rule were implemented, these representatives would need to gain additional certification in order to continue to serve their clients or attract new clients. The DOL cost-benefit analysis does not appear to take into account the significant costs that would be incurred by investment professionals in studying for and passing the certifying exam.

14. To illustrate the potential size of these costs, we understand that Primerica currently has 233 agents who hold a Series 65 license that would qualify them to provide advisory services if the Proposed Rule was implemented, in comparison with approximately 16,000 agents who do not currently hold that license. Based on what we believe is a conservative estimate of 50 hours of study and preparation time that would be required on average for an individual to prepare for the Series 65 exam, and valuing that time at the 2009 median hourly wage for personal financial advisers, \$32.79,[8] if 60% of Primerica's agents chose to become investment advisers after the implementation of the Proposed Rule, the additional cost incurred would be \$15.7 million (= 16,000 X 60% X 50 X \$32.79). Of course, this calculation is necessarily a rough approximation, but it does indicate that the cost of additional certifications alone could easily double, if not more than double, the DOL's estimate of the costs associated with the Proposed Rule, especially considering that this figure is based on the representatives of only one company among many in the industry. Industry-wide, we understand that there are more than 300,000 registered representatives in the U.S. which are not licensed to provide advisory services;

8. Bureau of Labor Statistics May 2009 National Occupational Employment and Wage Estimates, available online at http://www.bls.gov/oes/current/oes_nat.htm#13-0000.

therefore, the costs calculated above, extrapolated to the entire industry, would reach over $295 million.[9]

15. This estimate also does not include other potentially large economic losses associated with the higher certification requirements. First, faced with the costs of new certification, many brokerage representatives would likely choose not to acquire the necessary certification and therefore potentially leave the industry. Moreover, we understand that brokerage representatives are one of the only sources of financial information some investors encounter, so even aside from the other effects we discuss below, fewer professionals employed in the industry could lead to lower levels of retirement savings and lower (or even no) returns experienced by these investors.

B. Expenses Paid By IRA Investors Would Likely Rise Significantly Due to the Proposed Rule.

16. More importantly, the Proposed Rule would likely significantly increase the expenses paid by IRA investors for several reasons. First, under the Proposed Rule, firms that currently do not have fiduciary status with respect to their IRA customers would become fiduciaries. This would create additional costs for these firms to acquire and maintain new client accounts, due to the increased compliance costs associated with fiduciary status, and more importantly, the heightened risk of litigation faced by fiduciaries. Economic principles indicate that at least some of these costs would be passed on to investors in the form of higher prices to open and/or maintain IRA accounts. A basic principle of economics is that prices charged by firms in an industry will rise if firms face an increase in per-unit or per-client costs, such as that the increased costs associated with compliance and litigation risks.[10]

---

9. 300,000 X 60% X 50 X $32.79 = $295.1 million.

10. See, e.g., Robert E. Hall and Marc Lieberman, *Economics: Principles and Applications, 4th ed.*, at 251. The key exception to this principle would be in an industry in which consumer demand is infinitely elastic, for instance, if there is a perfect substitute for the good, such as black pens for blue pens. It is unlikely consumers perceive any perfect substitutes for financial services.

Therefore, it is likely that IRA investors will incur higher expenses under the Proposed Rule than they do now.

17. In addition, in the current regulatory environment, commission-based broker-dealer firms share the costs of opening and servicing IRA accounts with third-party providers of financial products. In practice, the way this sharing of costs occurs is that broker-dealer firms incur all of the costs up front, and then are partially recompensed by third-party product providers through commissions and other payments. We understand that this cost-sharing constitutes an important part of brokerage firms' business model.

18. Under the Proposed Rule, we understand that broker-dealers handling IRA accounts may be substantially restricted from receiving many forms of compensation from third-party product providers; in other words, they would no longer be able to share costs, and so would incur all of the costs of opening and servicing accounts themselves. Commission-based brokers would, in essence, face an increase in the cost of providing IRA investor services. As noted above, economic principles indicate that this increase in costs would likely cause prices to rise. At the same time, it is possible that, in the absence of commission payments, third-party product providers would reduce prices charged directly to investors, potentially offsetting to some degree the higher prices charged by brokers.[12] In connection with the Proposed Rule, the DOL has not presented any study of the overall impact that the elimination of this form of cost-sharing would have on total expenses paid by investors (nor are we aware of any conclusive evidence on this question from other sources), which is a key parameter necessary in order

12. While in theory, the decrease in costs faced by third-party product providers could be fully passed on to IRA investors, there are several industry-specific reasons why this effect would be unlikely to fully offset the increase in fees charged by financial services firms. First, investors who purchase through commission-based services represent only a fraction of total demand for these products, and costs faced by product providers from other distribution channels would not fall due to the Proposed Rule; in other words, to the extent that product providers experience decreases in costs due to the Proposed Rule, those savings would be spread out across all purchasers of the product, not only those who buy through commission-based services. Second, in the absence of payments to brokers, product providers may invest more in direct-to-consumer advertising or other methods of marketing.

to assess the full costs of the Proposed Rule. Nevertheless, this effect provides an additional reason, besides those mentioned above, why investors would likely face higher expenses in opening and/or maintaining IRA accounts under the Proposed Rule.

19. Available evidence is consistent with the premise that prices will be higher under the Proposed Rule. In the current regulatory environment, IRA investors can choose between broker-dealer firms offering commission-based service and certain other firms that provide "fee-only" service, in which advisers act as fiduciaries and forego most or all third-party commissions. While the fees any given investor pays usually depend somewhat on the details of his investments, as a general matter, a comparison of typical fees charged by the two types of service providers suggests that most IRA investors would incur higher expenses at fee-only firms than at commission-based firms, consistent with the notion that fiduciary status and the absence of third-party compensation result in higher expenses to investors.[13]

20. Primerica's fee structure, which we understand is typical for commission-based firms, charges investors a front load that is a percentage of assets purchased, with 5.5% being a typical rate, and then a custodial fee of $20 each year the account remains open.[14] By comparison, we understand that a typical fee-only adviser charges investors an annual fee calculated as a percentage of assets under management, with 1.5% being a typical rate, as well as an additional custodial fee similar to that charged by commission-based firms. While investors who make frequent trades or who have very short investment horizons may save through the use of fee-only services, investors who buy and hold

---

13. We understand that fee-based advisors typically provide additional services to investors not provided by commission-based brokers, and a full analysis would account for added value received by investors from these additional services; however, IRA investors, who primarily employ "buy-and-hold" strategies, typically have relatively few ongoing needs and so would be unlikely to benefit greatly from these services. Moreover, it is presumably the case that the additional benefits most commission-based account holders would receive with fee-based service would be lower than the additional expenses they would pay; otherwise, they would be employing fee-based advisers currently.

14. We understand that the size of the front load often declines when accounts reach a certain size.

investment assets for long periods, as IRA investors generally do, will typically pay lower fees with a commission-based service.[16] This is because, on a continuing basis, they pay only a low custodial fee every year instead of a percentage of an account which continues to grow in value (along with a similar-sized custodial fee).

21. For instance, consider an investor who opens a $2,000 IRA account invested in a typical balanced equity and mutual fund, and adds $100 per month to the account. Attachment 1 shows how the value of this investment, made in March 1991, would have grown over the following 20 years, with (a) a typical commission-based brokerage service charging a 5.5% front-load, and (b) a typical fee-only advisory service charging 1.5% annual fees.[17] In this example, the value of the investment would be higher under the commission-based service by June 1996, or in other words, so long as the investor held the IRA for more than five years and two months. As shown in the bottom panel of Attachment 1, by March 2011, the value of the investment under the commission-based expense schedule would be $10,931 higher than under the fee-only expense schedule.

22. Consistent with this analysis, a 2010 Oliver Wyman analysis performed for SIFMA found that, based on actual fees charged by 17 retail brokerage firms, typical investors would pay between 23 and 37 basis points more with fee-only accounts than with commission-based accounts under the current fee structure annually.[18]

---

16. The findings of the SEC's recent study of investment advisers and broker-dealers is consistent with this argument. See Securities Exchange Commission, "Study on Investment Advisers and Broker-Dealers", January 2011, at 152 (stating "[i]f, in response to the elimination of the broker-dealer exclusion, broker-dealers elected to convert their brokerage accounts from commission-based accounts to fee-based accounts, certain retail customers might face increased costs, and consequently the profitability of their investment decisions could be eroded, especially accounts that are not actively traded").

17. The underlying data from Morningstar used in this summary table is also attached. This analysis is based on an investment in Invesco Van Kampen Equity and Income Fund, Class A shares, and assumes all dividends and capital gains are re-invested in the fund. Since annual custodial fees are typically similar between commission-based and fee-only services, we ignore these here.

18. Oliver Wyman (2010), "Standard of Care Harmonization: Impact Assessment for SEC", October 2010.

23. The comparison between investor expenses under a commission-based brokerage service and expenses under a fee-only adviser service is analogous to a widely-accepted difference between investor expenses associated with the purchase of "class A" and "class C" mutual fund shares. Many funds offer multiple classes of shares, which differ only in the structure of the fund's expenses charged to investors. Typically, with "class A" shares, funds charge investors a front load fee, but then low annual fees on a continuing basis, while with "class C" shares, funds charge little or no front load, but higher annual expenses while holding the fund. It is widely noted that long-term "buy-and-hold" investors, which includes most IRA investors, pay lower fees by purchasing class A shares than they would with class C shares.[20] Analogously, most IRA investors pay lower fees with commission-based brokerage services than with fee-based advisory services.

24. Since IRAs are one of the primary means by which Americans save for retirement (38% of those saving for retirement hold IRAs)[21], even a small increase in fees on these accounts would impact a significant number of investors and lead to a large increase in costs in aggregate. In 2009, there was $4.2 trillion held in IRAs.[22] We understand that, based on the industry data analyzed by Oliver Wyman, they concluded that 66% of even the largest IRAs (those with more than $250,000 in assets) are held in accounts with commission-based brokerages, with this share much higher for smaller IRAs.[23] Even applying the low 66% figure to all IRA assets, if the Proposed Rule led to even a 1 basis point increase in annual costs relative to assets for these investors, it would generate $277 million (= 4.2 trillion X 66% X 0.01%) in additional expenses for investors annually, or over $2 billion over 10 years in current dollars,

---

20. See, e.g., FINRA, "Investor Alert: Understanding Mutual Fund Classes", Oct. 6, 2008, available at http://www.finra.org/Investors/ProtectYourself/InvestorAlerts/MutualFunds/p006022 (stating that for purchasers of class C shares, "in most cases, your total cost would be higher than with Class A shares, and even class B shares, if you hold for a long time"). See also Brian K. Reid and John D. Rea, "Mutual Fund Distribution Channels and Distribution Costs", *Perspective* 9(3), July 2003, at 13 (stating, "... investors subject to the maximum front-end sales load would prefer C shares for short and intermediate holding periods. Investors with a long investment horizon would choose A shares").

21. AARP, "AARP Bulletin Survey on Retirement Savings: Executive Summary", April 2009, at 4.

22. Investment Company Institute, "The U.S. Retirement Market, 2009", at 2.

23. Oliver Wyman (2011), *supra*, at 11.

using a discount rate of 7%. This illustrative calculation clearly indicates the potential for much higher costs from the Proposed Rule than estimated by DOL.

C. *Returns on Investors' Retirement Funds Would Likely Decline Due to the Proposed Rule.*

25. For several reasons, under the Proposed Rule, investors would likely reduce their usage of IRAs, as well as brokerage services associated with IRAs. First, because, as described above, prices for financial services would likely rise, the Proposed Rule would be likely to cause some individuals to choose not to open IRA accounts or to invest less in them. In addition, under the Proposed Rule, investors would likely face higher minimum account balance requirements to open an IRA. Firms impose minimum account balance requirements because for relatively small accounts, the cost incurred by a firm in opening or servicing the account may be higher than the revenue received. As discussed above, under the Proposed Rule, per-account costs incurred by commission-based brokerage firms would likely increase. Because the revenue generated by low balance accounts is small, an increase in costs would likely mean that these firms would increase minimum account balance requirements for IRA investors.

26. Available evidence is consistent with increases in minimum required account balances under the Proposed Rule. Minimum account sizes generally appear to be substantially higher among fee-based advisers, who incur the expense of fiduciary status and forego most third-party compensation, than among commission-based advisers. Primerica, for instance, allows investors to open IRA accounts with as little as $250, while we understand that minimum account sizes for fee-based advisers are typically more than $10,000, and often $50,000 or more. We have been informed that Primerica believes it would need to raise its minimum IRA account size to around $25,000 if it were forced by the Proposed Rule to forego third-party commissions.

27. Small investors constitute the bulk of IRA investors. We understand that, based on the industry data analyzed by Oliver Wyman they concluded that 51% of IRA accounts include less than $25,000 in assets, and more than 30% of IRA accounts have asset values below the current minimum balance requirement for fee-based advisory services at *any* of the firms providing data for their sample.[24] Unless advisory firms substantially reduced these minimum requirements, investors would need to either move their funds to self-directed brokerage accounts and forego the services they currently receive, or else move their funds out of IRAs altogether. In some cases, investors could add funds to their accounts to reach the new, higher, minimum account balances, although because the IRS limits annual contributions to IRAs, only investors who hold amounts relatively close the new minimum account balance requirements would have this option.[25] Therefore, the Proposed Rule could lead to a reduction in the rate at which individuals invest in IRAs and receive financial services in connection with IRAs.

28. Besides forcing changes for some current IRA investors, increases in minimum account balances would also impact investors seeking to open new IRAs, particularly since new accounts often start with low balances. Because, in the absence of the Proposed Rule, these accounts would be expected to grow over time, the long-run impact in reducing the amount of funds held in IRAs could be even larger than the immediate impact.

29. Consistent with the implications of higher prices and lower investment participation from the Proposed Rule, analysts predict that a similar regulation reducing sales commissions paid to investment advisers in the United Kingdom will raise fees to investors and lead to a dramatic reduction in the size of the financial adviser industry. As reported by analysts at Ernst & Young, under the new

---

24. Oliver Wyman (2011), *supra*, at 10 and 17 (showing 22 million total accounts analyzed and 7.2 million accounts with insufficient assets to access the advisory channel at any firm).

25. For 2010, the IRS limits IRA contributions below $5,000 for individuals under age 50, and $6,000 for individuals above age 50. See Internal Revenue Service, *Publication 590: Individual Retirement Accounts (IRAs)*, at 6.

rules, "[f]irst, it seems likely that the mass market and the typical bank customer will not be enthusiastic about paying the sort of fees that make offering the advice attractive. Second, simplified advice becomes a major economic challenge, requiring a radically reduced cost base if it is to present a solution for the mass market ... There is a real possibility that the independent advisory sector, as we know it, will shrink significantly."[28]

30. As individuals reduce their holdings of assets in IRAs, they may choose to invest in other, less tax-privileged vehicles, they may choose to invest without the financial services they previously employed, or they may simply choose to invest less overall. Any of these effects would lead to lower investment returns to individual investors, exacerbating the prevailing retirement savings shortages in the U.S.[29]

31. As noted above, IRA investments in the U.S. totaled $4.2 trillion in 2009, constituting 26% of all retirement assets.[30] If, because of the effects described above, the Proposed Rule led even 1% of investment assets to be withdrawn from IRA accounts, and if those assets therefore generated 25 basis points lower annual returns (due, e.g., to disadvantaged tax treatment of non-IRA funds or because of the absence of broker services), then the Proposed Rule would generate losses of nearly $790 million in current dollars over ten years using a 7% discount rate. This would dramatically increase the costs of the Proposed Rule far beyond the level anticipated by the DOL. In addition, investors who withdraw funds from IRAs typically pay tax penalties for early withdrawal, further adding to the costs of the Proposed Rule.

---

28. Ernst & Young, "RaDaR: Life and Pensions Outlook for 2011", January 2011, at 7 and 9.

29. According to a 2009 study, the average American family faces a 37% shortfall in the income they will need for retirement. (See McKinsey & Co., "Restoring Americans' Retirement Security: A Shared Responsibility", 2009, at 2.)

30. Investment Company Institute, "The U.S. Retirement Market, 2009", at 2.

III. THE EVIDENCE PRESENTED BY THE DOL SUPPORTING ALLEGED BENEFITS FROM THE PROPOSED RULE DOES NOT PROVIDE A SUFFICIENT BASIS TO CONCLUDE THAT THESE BENEFITS WOULD BE LARGE ENOUGH TO OUTWEIGH THE COSTS.

32. The DOL cost-benefit analysis claims three specific benefits would follow from implementation of the Proposed Rule: (a) discouraging harmful conflicts of interest in which "... service providers strike deals that profit one another at the plan's expense or subordinate the plan's interest to someone else"[31]; (b) providing pension plans with "better value for the service fees they pay", along with "the ancillary benefit of improved returns on plan assets"[32]; and (c) enhancing "the Department's ability to redress service provider abuses that currently exist in the market"[33].

33. We understand that IRAs are outside the scope of the DOL's enforcement authority. Therefore, the potential benefit from the Proposed Rule in enhancing the effectiveness of DOL's enforcement initiatives, while potentially relevant in other segments of the financial services industry, would not by itself provide a basis for supporting an extension of fiduciary status to broker-dealers providing services to IRA investors. For that reason, we will focus our attention on evaluating the evidence presented by the DOL to support the other two claimed benefits.

34. The only quantification of benefits provided in the DOL cost-benefit analysis is the following statement: "If just 10 percent of plans realize a one basis point (0.01 percent of plan assets) service value improvement, it would be worth approximately $399 million over ten years ..."[34]. In preparing this Comment, we reviewed the literature cited by the DOL, as well as the broader economic and financial literature related to these issues, and did not find any statistical study which concludes that the Proposed Rule or any similar regulation would generate a one basis point improvement (or any size improvement) in pension plan service value.

---

31. DOL cost-benefit analysis, *supra*, at 65272, section 5(a).
32. DOL cost-benefit analysis, *supra*, at 65273, section 5(b).
33. DOL cost-benefit analysis, *supra*, at 65273, section 5(c).
34. DOL cost-benefit analysis, *supra*, at 65273, section 5(b).

*A.	The Cited Studies Do Not Provide a Basis to Conclude that Benefits from the Proposed Rule Would Be Large Enough to Outweigh the Costs.*

35. The DOL argues, qualitatively, that benefits of some size may accrue from the Proposed Rule by citing four statistical studies that the DOL claims support the hypothesis that potential conflicts of interest faced by financial service providers harm investors. Specifically, the DOL cites a study performed by the U.S. Government Accountability Office ("GAO") and three other studies presented in unpublished academic manuscripts. A review of these four studies indicates that none of them focus specifically on IRA investments, nor does any claim to provide direct evidence regarding the Proposed Rule or any similar rule.

36. As we discussed above, the Proposed Rule would likely change the market for IRA investment services significantly. However, each of the studies cited by the DOL analyze investor behavior in the current regulatory environment and under the current industry structure. Therefore, even interpreting the findings of these studies in ways favorable to the DOL's claims, these findings would not be sufficient to conclude that a major change in the regulatory environment, as under the Proposed Rule, would lead to significant benefits because the Proposed Rule could change the structure of the industry in a variety of ways not considered by these studies. In order to understand the full impact of the Proposed Rule, the DOL would need to study carefully how the Proposed Rule would impact investor behavior under the significantly changed industry that the Proposed Rule would impose.

37. In any case, the cited studies do not in fact provide strong evidence that potential conflicts of interest faced by investment advisers lead to significant reductions in investor value, as some of the authors of the studies themselves indicate. Moreover, the results of these studies are consistent with significant benefits received by investors from financial service providers in the current regulatory environment. We first discuss the referenced GAO study and then the three referenced academic studies.

(i) The cited GAO study does not provide a basis to conclude that the benefits from the Proposed Rule would be large enough to outweigh the costs.

38. According to the DOL, the GAO study "links undisclosed conflicts with 130 basis points of underperformance in defined benefit pension plans".[35] Clearly, then, this study's results would not directly apply to IRA investment advisers, since IRAs are not defined benefit plans.[36] Defined benefit plans typically cover aggregate retirement benefits for most or all employees of a large corporation or other institution, while defined contribution plans like IRAs are held by individual investors saving for their own retirement. The advisory needs of defined benefit plan managers are therefore very different from those of IRA investors. Moreover, as recognized in the referenced quote from the DOL, the GAO study would only show, at most, that *disclosure* of potential conflicts of interest could lead to the claimed benefits, not that eliminating the source of these potential conflicts is necessary to achieve such benefits.

39. Moreover, the authors of the GAO study, as well as the individuals who collected the data employed in the GAO study, clearly indicate that their results cannot support a conclusion that conflicts of interest generate investor harms. The first paragraph of the GAO study indicates, "[b]ecause many factors can affect returns, and data as well as modeling limitations limit the ability to generalize and interpret the results, this finding should not be considered proof of causality between conflicts and

35. DOL cost-benefit analysis, *supra*, at 65272, section 5(a). The GAO study is described in two separate documents, which we will refer to interchangeably as "the GAO study", but separately as "GAO (2009)" and "GAO (2007)": (1) congressional testimony summarizing the study (GAO, "Conflicts of Interest Can Affect Defined Benefit and Defined Contribution Plans", GAO-09-503T, March 24, 2009); and (2) the details of the statistical analysis performed by GAO (GAO, "Conflicts of Interest Involving High Risk or Terminated Plans Pose Enforcement Challenges", GAO-07-703, June 2007).

36. In congressional testimony summarizing the study (GAO 2009), the GAO's Acting Director discussed the possibility that conflicts of interest could also affect defined contribution plans, and cited some evidence that defined contribution plan sponsors and participants may not be fully aware of potential conflicts of interest faced by their pension consultants, but he nevertheless emphasized that the GAO only specifically studied the impact of conflicts of interest on defined benefit plans, noting "[o]ur study focused exclusively on DB [defined benefit] plans and less information exists on the extent of nature of conflicts of interest in the DC [defined contribution] plan environment" (GAO 2009, *supra*, at summary page).

lower rates of return".[37] The GAO also recognizes that, even if harm due to conflicts of interest could be shown with respect to the specific pension plan consultants they studied, such results would not imply harm to retirement investors generally: "… these results cannot be generalized to the population of pension consultants since the consultants examined by the SEC were not selected randomly",[38] and "the plans included in the analysis should not be considered as representative of the population of defined benefit pension plans".[39]

40. As suggested by the quotation above, the sample of pension plan service providers studied by the GAO was collected, and potential conflicts of interest identified, in an earlier analysis by the Securities and Exchange Commission ("SEC") in 2002 and 2003. In their analysis of these providers' behavior, however, the SEC did not conclude that any of them actually *acted* on undisclosed potential conflicts of interest to the detriment of clients, in fact stating that "[w]e could not fully analyze whether pension consultants 'skewed' their recommendations to favor certain money managers"[40]. As noted above, nothing in the SEC's study showed that firms providing financial services to IRA investors had, or acted on, conflicts of interest, because none of the plans analyzed by the SEC were IRAs, or defined contribution plans of any sort.

(ii) Academic studies cited by DOL do not support a conclusion that the alleged benefits from the Proposed Rule would be large enough to outweigh the costs.

41. Turning next to the academic studies cited in the DOL cost-benefit analysis, the results presented in these papers also provide little evidence supporting the DOL's claimed benefits, and in fact are generally consistent with significant benefits accruing to investors from financial services in the current regulatory environment. The first two of the cited studies are similar. Bergstresser, et al. (2007) and Bullard, et al. (2007) both compare returns achieved by "no-load" mutual funds, which are

37. GAO (2009), *supra*, at summary page.
38. GAO (2007), *supra*, at 16.
39. GAO (2007), *supra*, at 43.
40. SEC (2005) "Staff Report Concerning Examinations of Select Pension Consultants", Office of Compliance Inspections and Examinations, May 16, 2005, at 5.

commonly sold directly to investors, and "load" mutual funds, which are more commonly sold through brokers or other intermediaries.[41] Bergstresser, et al. (2007) argue that load funds underperform, on average, no-load funds. Bullard, et al. (2007) argue that investors in both load and no-load funds tend to mis-time market transactions and so underperform a simple "buy-and-hold" strategy, but that investors in load funds underperform by a greater degree.

42. A key issue in interpreting these results is potential differences between individuals who make use of financial services and those who do not. For instance, investors who make use of these services may make poorer financial decisions than others who do not use these services. Providing support and information to individuals who might otherwise make poor decisions is, after all, the purpose of these services. If investors who make use of these services have poorer financial decision-making skills, a broker or adviser may provide value by supporting better decisions, even if they cannot fully eliminate investors' tendencies to make poor financial decisions. In this case, investors who make use of financial services may still achieve lower returns than those who do not, yet they would earn even lower returns in the absence of the broker or adviser's help.

43. This issue is explicitly recognized by Bullard, et al. (2007), who indicate that their results could be explained by the "well-documented psychological tendency of investors to overweight recent performance. Although investment professionals presumably are more aware of, and less susceptible to, a short-term performance bias, their clients might be more susceptible to this bias than self-directed investors. Those who seek out professional guidance may be less knowledgeable about investing and more inclined to expect or pressure their advisors to trade on short-term performance."[42]

---

41. Daniel Bergstresser, John Chalmers, and Peter Tufano (2007) "Assessing the Costs and Benefits of Brokers in the Mutual Fund Industry", Social Science Research Network Abstract 616981, Sept. 2007; Mercer Bullard, Geoff Friesen, and Travis Sapp (2007) "Investor Timing and Fund Distribution Channels", Social Science Research Network Abstract 1070545, Dec. 2007. Begstresser, et al. (2007) has since been published in an academic journal (*The Review of Financial Studies*) in a form similar to the earlier working paper; because the DOL cost-benefit analysis relied on the working paper, we will refer to that version of the manuscript here as well.

42. Bullard, et al. (2007), *supra*, at 11.

44. Bergstresser, et al. (2007) also recognize that their results do not prove that conflicts of interest cause investor losses, and in fact they state that their results could also be evidence of significant benefits to investors from investment brokers, writing "[o]ur results are consistent either with substantial non-tangible benefits delivered by the broker-distributed sector or with conflicts of interest between brokers and their clients."[43]

45. Unlike the two academic studies discussed above, the third study cited in the DOL cost-benefit analysis, Zhao (2005), does not examine investment returns, but focuses instead on which funds investors purchase.[44] Specifically, the author compares the inflow of investor dollars into load funds with different fees, finding that funds with higher fees receive greater inflows of investor dollars on average. In the absence of specific information on how fees for the funds studied by Zhao translate into compensation for brokers or advisers, this result does not provide strong evidence on the specific impact of conflicts of interest.

46. Moreover, this result may be equally consistent with the result predicted by basic economic theory in the absence of conflicts of interest or other market failures. Specifically, Zhao's result that more popular funds, which receive larger inflows, charge higher loads to investors, is not necessarily surprising, since this is an effect that is common to popular products of all types. The fact that more expensive restaurants are often more popular than less expensive restaurants does not necessarily imply that restaurants are exploiting diner ignorance of meal prices; instead, more popular restaurants can simply charge higher prices because the demand for their meals is higher. Zhao's study cannot determine whether investors are purchasing certain funds because opportunistic brokers are pushing investors to buy those funds, or whether investors simply prefer those funds for some other reason.

---

43. Bergstresser, et al. (2007), *supra*, at 1.
44. Xinge Zhao (2005) "The Role of Brokers and Financial Advisors Behind Investments into Load Funds", China Europe International Business School Working Paper, Dec. 2005.

47. These three academic studies also uncovered evidence that financial services firms provide significant benefits to investors. Bergstresser, et al. (2007) found supporting evidence for the hypothesis that "brokers specialize in unique products, especially ones that individual investors would have difficulty in evaluating on their own",[45] and they further acknowledge that "[b]rokers may help their clients save more than they would otherwise save, they may help clients more efficiently use scarce time, they may help customize portfolios to investors' risk tolerances, and they may increase overall investor comfort with their investment decisions"[46]. Zhao (2005) similarly found evidence consistent with the notion that investment advisers provide valuable services to their clients, stating "... while no load fund investors flock into larger funds with more visibility, load fund investors are more likely to be directed by brokers and financial advisers into smaller funds, which might experience better performance than larger funds exceeding their optimal size."[47] These studies do not attempt to quantify the value of these benefits, which may be large.

48. Moreover, these studies do not attempt to calculate other value provided by brokerage services unrelated to the returns achieved by plan assets, such as the provision of information regarding the appropriate amount of current investment needed to meet retirement goals or the tax consequences of taking a distribution from an IRA after a rollover event.

IV. A REVIEW OF ECONOMIC THEORY AND AVAILABLE EVIDENCE REGARDING THE IRA INVESTMENT SERVICES INDUSTRY DOES NOT SUPPORT A CONCLUSION THAT THE PROPOSED RULE WOULD GENERATE BENEFITS LARGE ENOUGH TO OUTWEIGH THE COSTS.

49. It may be true that certain brokers providing services to investors face *potential* conflicts of interest; nevertheless, there are two key economic factors that economic theory and available evidence indicate constrain firms from *acting* on these conflicts to the detriment of customers:

---

45. Bergstresser, et al. (2007), *supra*, at 13.
46. Bergstresser, et al. (2007), *supra*, at 2-3.
47. Zhao (2005), *supra*, at 32.

competition and reputational concerns. In addition, we understand that broker-dealers currently face a range of regulations that further diminish the impact of potential conflicts of interest. Because these constraints already exist in the current market environment, we see no basis to conclude that implementation of the Proposed Rule would produce significant additional benefits beyond what investors currently receive. Therefore whatever benefits may result from the Proposed Rule would be unlikely to be large enough to outweigh the costs we identified above.

50. The revealed preference of investors themselves for commission-based investment services with respect to IRAs provides perhaps the most powerful evidence that, even in the presence of potential incentives for broker-dealers to behave opportunistically with respect to their clients, market discipline protects investors. We understand that, based on the industry data analyzed by Oliver Wyman, they concluded that 88% of all IRA accounts are held with commission-based brokerage firms.[48] According to the Investment Company Institute, "80 percent of [households] that owned funds outside a workplace retirement plan held funds purchased through a professional adviser."[49] Moreover, "[h]alf of all mutual fund shareholders indicated they had ongoing relationships with financial advisers",[50] illustrating the continuing value investors perceive in their relationships with financial service providers.

51. Basic economic principles indicate that competition among firms places a constraint on the ability of these firms to behave in ways detrimental to their clients because, to the extent they do so, their clients will experience low benefits from the services provided and create opportunities for competing firms, which can earn sales by pointing out the low benefits a customer is currently receiving and offering to provide higher benefits. In addition, competition in the market for expert advice creates opportunities for investors to readily seek "second opinions", which provide an additional check on advisers' ability to exploit informational advantages they may have with respect to their customers.

---

48. Oliver Wyman (2011),*supra*, at 11.
49. Investment Company Institute, *2010 Investment Company Fact Book*, at 85.
50. Investment Company Institute, *2010 Investment Company Fact Book*, at 86.

52. In fact, the market for IRA investment services appears to be highly competitive, as a wide range of evidence indicates.

- Recent industry reports indicate that there are around 25,000 companies in the U.S. competing to provide financial planning and investment advice to individuals and businesses.[51]
- Public financial filings by companies in the industry consistently indicate a high degree of competition: "We operate in a highly competitive environment with respect to the sale of financial products" (Primerica)[52]; "We operate in a highly competitive industry" (Ameriprise Financial)[53]; "We are subject to competition in all aspects of our business" (LPL)[54]; "All aspects of the Partnership's business are highly competitive" (Edward Jones)[55].
- Average fees for mutual funds, one of the major products purchased through commission-based brokers, have declined consistently and dramatically over time. As noted by the Investment Company Institute in 2010, "[f]ees and expenses incurred by stock and bond mutual fund investors have declined by half since 1990".[56] Declining prices are typical of competitive industries, and accordingly, this industry publication attributes the noted decline in fees at least partially to competition.[57]
- Investors are increasingly moving towards lower-cost mutual funds and other investments, providing incentives for brokers to keep the fees they charge low in order to remain competitive. One way to see investors' competitive pressure with respect to costs is to compare the average expense ratio on all mutual funds offered in the marketplace with the average

---

51. First Research "Industry Profile: Financial Planners and Investment Advisors", Oct. 26, 2009.
52. Primerica Inc. 10-K for the fiscal year ended December 31, 2010, at 35.
53. Ameriprise Financial, Inc. 10-K for the fiscal year ended December 31, 2010, at 17.
54. LPL Investment Holdings Inc. 10-K for the fiscal year ended December 31, 2010, at 19.
55. The Jones Financial Companies LLLP 10-K for the fiscal year ended December 31, 2010, at 18.
56. Investment Company Institute, *2010 Investment Company Fact Book*, at 64. These findings are broadly consistent with trends in expense ratios studied previously by the GAO ("Information on Trends in Fees and Their Related Disclosure", GAO-03-551T, Testimony Before the Subcommittee on Capital Markets, Insurance and Government Sponsored Enterprises, Committee on Financial Services, House of Representatives, March 12, 2003, at summary page).
57. Investment Company Institute, *2010 Investment Company Fact Book*, at 66.

expense ratio actually paid by investors. To the extent that shareholders invest more in lower-cost funds, they will pay lower expenses than charged by the average fund. In 1995, the average fund charged an expense ratio of 1.52%, while investors actually paid an average expense ratio of 1.04%, a difference of 48 basis points.[58] By 2009, the average fund charged an expense ratio of 1.50%, while investors paid an average of 0.86%, a difference of 64 basis points.

53. Several academic studies of markets for expert advice support the hypothesis that competition is effective in constraining conflicts of interest specifically in the market for expert advisers:

- Bolton, et al. (2005) developed a model of the provision of information by sellers of financial services to customers, as takes place in the IRA investment broker industry. They conclude that "competition both reduces the gains from lying and induces financial institutions to disclose information".[59] They further write that their results "... directly challenge the conventional wisdom that information is only credible if it is provided by an independent institution that has no such conflicts of interest."[60]
- Krausz and Paroush (2002) analyze conflicts of interest in financial advising, and argue that "... competition reduces transactions cost and it is easier for dissatisfied investors to transfer from one financial advisor to another. Furthermore, if deception is very severe, competition from other financial advisors and institutions will erode the financial advisor's returns, yet again reducing the incentive to deceive."[61]

58. Investment Company Institute, *2010 Investment Company Fact Book*, at 66.
59. Patrick Bolton, Xavier Freixas, and Joel Shapiro (2007) "Conflicts of Interest, Information Provision, and Competition in the Financial Services Industry", *Journal of Financial Economics* 85(2), at 298.
60. Bolton, et al. (2007), *supra*, at 298.
61. Miriam Krausz and Jacob Paroush, "Financial Advising in the Presence of Conflict of Interests", 54 *Journal of Economics and Business*, at 57.

- Patron and Roskelley (2007) analyze the analogous case of conflicts of interest in markets for expert real estate advice, and conclude that "agents are less likely to suggest aggressive bargaining strategies [for their clients] when there is little market competition."[62]

As on most topics, there are a wide variety of different opinions expressed by academic authors depending on the assumptions they make and the methodologies they employ; however, at the very least these articles indicate that, without further detailed study of the IRA investment brokerage industry, it is highly premature to conclude that the Proposed Rule will result in significant benefits to investors or others large enough to outweigh its costs.

54. A second key factor that constrains investment brokers from acting on potential conflicts of interest to the detriment of their clients is the need to maintain a positive reputation among clients and potential clients. Economic principles indicate that if a firm develops a reputation for low-quality service, its clients will be less likely to use that firm's services in the future, and will be less willing to recommend the firm's services to others. In the context of financial services firms, this provides an incentive for those firms to provide high-quality service to their clients, even in the presence of potential conflicts of interest.

55. Reputation as a factor limiting opportunistic behavior by firms has been repeatedly identified as important in the academic literature. In an extensive survey, MacLeod (2007) summarizes the substantial literature on this issue, which he describes as based on the premise that, "in a free market, sellers of high quality goods treat their reputation as an asset that loses its value should they choose to supply goods of low quality",[63] and that "reputation is an asset whose value is destroyed

62. Hilde E. Patron and Kenneth D. Roskelley, "The Effect of Reputation and Competition on the Advice of Real Estate Agents", 37 *Journal of Real Estate Financial Economics*, at 387.

63. W. Bentley MacLeod (2007) "Reputations, Relationships, and Contract Enforcement", *Journal of Economic Literature* XLV:595-628, at 596.

when a seller or buyer breaches their obligation."[64] Other standard and well-cited economic articles based on this concept are Klein and Leffler (1981)[65] and Rogerson (1983)[66].

56. Available evidence also indicates that reputation is an important factor specifically in the market for IRA investment services.

- Public financial filings by companies in the industry indicate that reputation is a key element in their business success: "Our reputation is one of our most important assets ... Damage to our reputation could cause significant harm to our business and prospects ... Our reputation is also dependent on our continued identification of and mitigation against conflicts of interest ... our reputation could be damaged if we fail, or appear to fail, to deal appropriately with conflicts of interest" (Ameriprise Financial)[67]; "We have spent many years developing our reputation for integrity and superior client service ... Damage to our reputation could cause significant harm to our business and prospects" (LPL)[68].
- Firms in the IRA investment services industry rely to a significant degree on referrals from clients to attract new business.[69] Referrals are only an effective source of sales leads when a firm holds a good reputation with its client base.

57. In summary, economic theory and available evidence indicate that factors in the current market environment likely serve to substantially constrain the ability of IRA investment brokers to act on

---

64. MacLeod (2007), *supra*, at 603.
65. Benjamin Klein and Keith B. Leffler (1981) "The Role of Market Forces in Assuring Contractual Performance", *Journal of Political Economy* 89(4):615-41.
66. William P. Rogerson (1983) "Reputation and Product Quality", *Bell Journal of Economics* 14(2):508-16.
67. Ameriprise Financial, Inc. 10-K for the fiscal year ended December 31, 2010, at 47.
68. LPL Investment Holdings Inc. 10-K for the fiscal year ended December 31, 2010, at 29.
69. First Research, "Industry Profile: Financial Planners and Investment Advisors", Oct. 26, 2009 (stating, "[i]nvestment advisers depend heavily on referrals for new customers"); see also Dow Jones News Service, "Wealth Adviser: Facing the Competition – Whatever it May Be", June 8, 2010 (stating, "[i]n the end, many advisers find the most productive approach is to focus on doing the best they can for the clients they have and counting on that to bring in referrals").

potential conflicts of interest. Therefore, we see no basis to conclude that the Proposed Rule would generate benefits large enough to outweigh the costs.

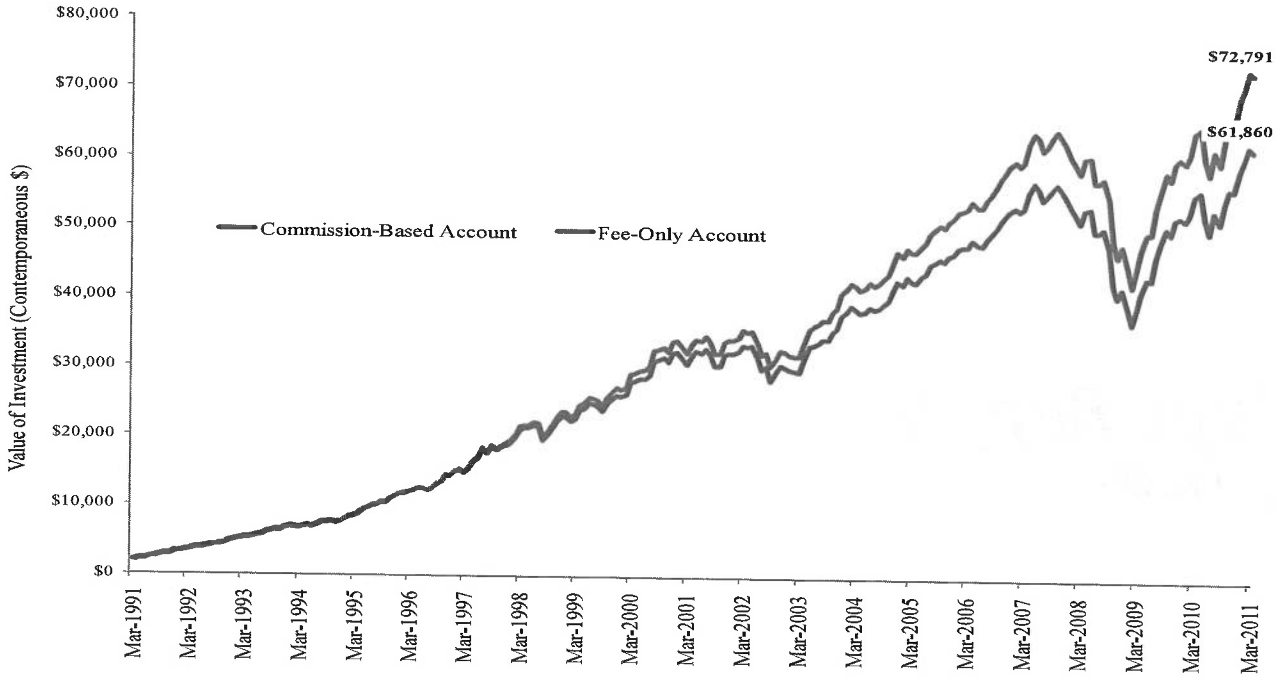
Daniel R. Fischel

Todd D. Kendall

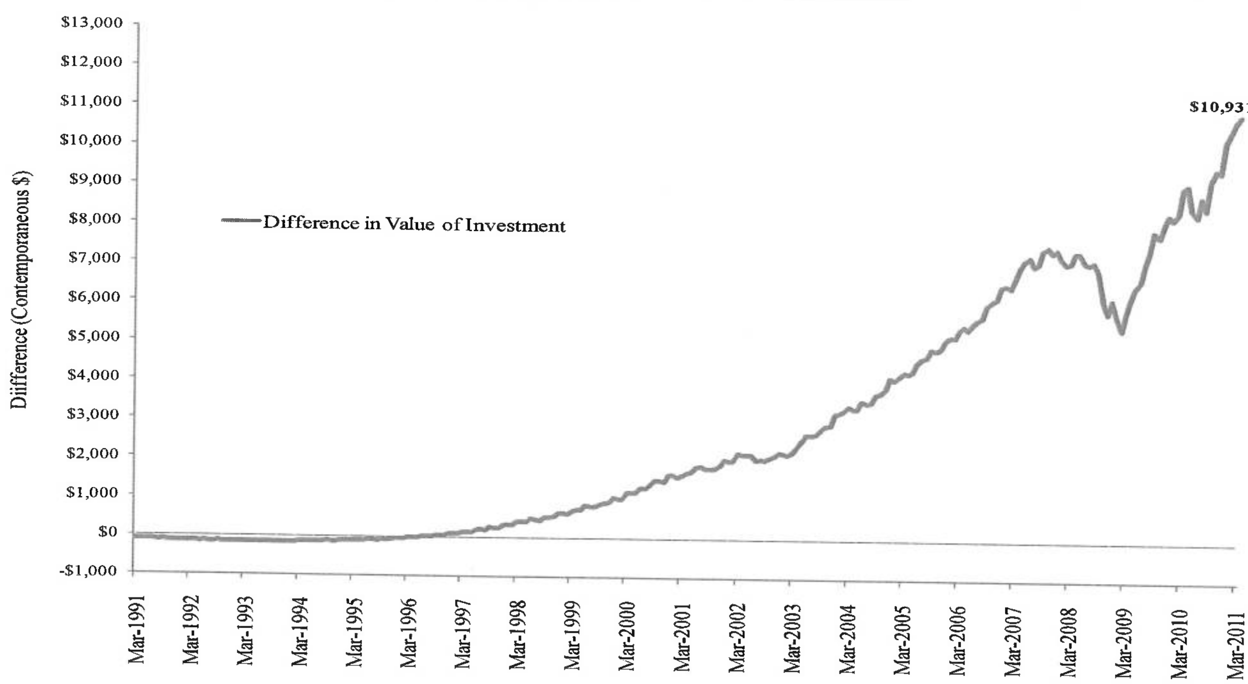
ATTACHMENT 1

Value of IRA Mutual Fund Investment Under Commission-Based and Fee-Only Expense Schedule ($2,000 Initial Investment and $100 Monthly Contributions)

March 31, 1991 through March 31, 2011



Difference in Value of Investment



Source: Morningstar analyses of Invesco Van Kampen Equity and Income Fund, Class A shares.

Notes: Commission-based expense schedule includes 5.5% front load on initial investment and contributions until asset value reaches $50,000, when front load on additional contributions declines.
Fee-only expense schedule includes 1.5% annual fee charged on assets under management.
Investment value assumes all dividends and capital gains are reinvested in the specified fund.
Value calculations do not incorporate taxes.

PFS INVESTMENTS INC.

Member of FINRA

Hypo Report

April 10, 2011

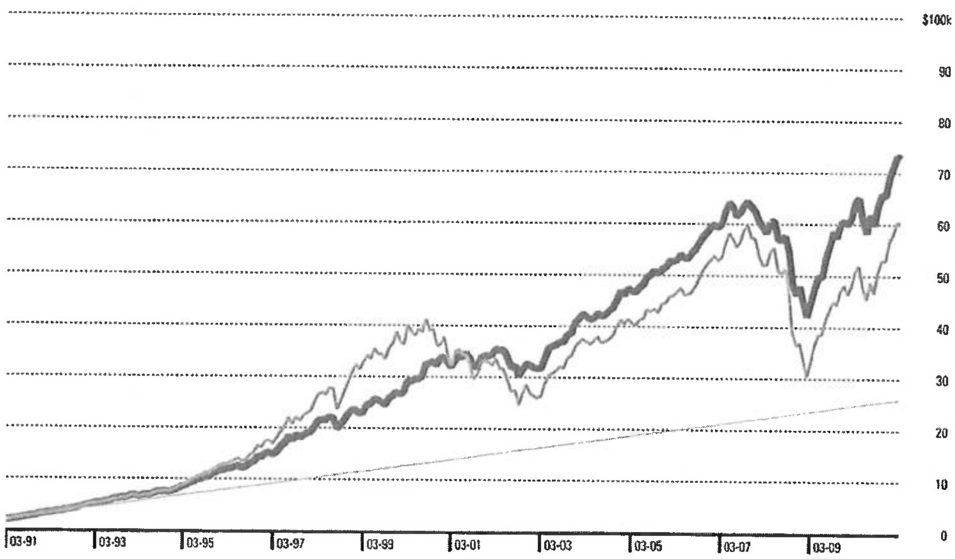
Prepared Especially For:

Your Representative:

Hypothetical Portfolio Illustration

03-31-1991 to 03-31-2011

Portfolio Summary



Planning Assumptions

Currency	USD
Rebalance	None
Federal Income Tax Rate	0%
Capital Gain Tax Rate	0%
State Tax Rate	0%
Tax Paid	Out of Pocket

Performance

Net Amount Invested	$26,000
Final Market Value	$72,791
Average Annualized Return	8.75%
Cumulative Return	435.73%

Investment Detail

Period		Beginning Balance	New Investment	Distribution/ Withdrawal	Total Reinvest	Charges & Fees	Taxes Due	Market Value	Total Return %
Totals		0	26,000	0	34,210	1,371	0	72,791	8.75
March	1991	0	2,000	0	0	110	0	1,890	-5.50
April	1991	1,890	100	0	0	6	0	1,989	-0.06
May	1991	1,989	100	0	0	6	0	2,147	2.93
June	1991	2,147	100	0	24	6	0	2,174	-3.38
July	1991	2,174	100	0	0	6	0	2,360	3.91
August	1991	2,360	100	0	0	6	0	2,517	2.43
September	1991	2,517	100	0	27	6	0	2,612	-0.21
October	1991	2,612	100	0	0	6	0	2,752	1.54
November	1991	2,752	100	0	0	6	0	2,781	-2.57
December	1991	2,781	100	0	40	6	0	3,096	7.71
January	1992	3,096	100	0	0	6	0	3,171	-0.80
February	1992	3,171	100	0	0	6	0	3,318	1.49
March	1992	3,318	100	0	34	6	0	3,358	-1.81
April	1992	3,358	100	0	0	6	0	3,516	1.73
May	1992	3,516	100	0	0	6	0	3,662	1.29
June	1992	3,662	100	0	37	6	0	3,719	-1.17
July	1992	3,719	100	0	0	6	0	3,937	3.18
August	1992	3,937	100	0	0	6	0	4,000	-0.94
September	1992	4,000	100	0	41	6	0	4,143	1.08
October	1992	4,143	100	0	0	6	0	4,254	0.27
November	1992	4,254	100	0	0	6	0	4,460	2.49
December	1992	4,460	100	0	39	6	0	4,639	1.76
January	1993	4,639	100	0	0	6	0	4,841	2.22
February	1993	4,841	100	0	0	6	0	5,009	1.41

Hypothetical Portfolio Illustration

03-31-1991 to 03-31-2011

Investment Detail

Period		Beginning Balance	New Investment	Distribution/ Withdrawal	Total Reinvest	Charges & Fees	Taxes Due	Market Value	Total Return %
Totals		0	26,000	0	34,210	1,371	0	72,791	8.75
March	1993	5,009	100	0	70	6	0	5,231	2.42
April	1993	5,231	100	0	0	6	0	5,238	-1.77
May	1993	5,238	100	0	0	6	0	5,402	1.21
June	1993	5,402	100	0	42	6	0	5,579	1.44
July	1993	5,579	100	0	0	6	0	5,725	0.82
August	1993	5,725	100	0	0	6	0	6,008	3.19
September	1993	6,008	100	0	45	6	0	6,158	0.83
October	1993	6,158	100	0	0	6	0	6,329	1.15
November	1993	6,329	100	0	0	6	0	6,324	-1.66
December	1993	6,324	100	0	266	6	0	6,588	2.60
January	1994	6,588	100	0	0	6	0	6,885	2.98
February	1994	6,885	100	0	0	6	0	6,799	-2.70
March	1994	6,799	100	0	97	6	0	6,632	-3.92
April	1994	6,632	100	0	0	6	0	6,790	0.86
May	1994	6,790	100	0	0	6	0	6,922	0.48
June	1994	6,922	100	0	54	6	0	6,889	-1.93
July	1994	6,889	100	0	0	6	0	7,168	2.60
August	1994	7,168	100	0	0	6	0	7,476	2.91
September	1994	7,476	100	0	57	6	0	7,437	-1.86
October	1994	7,437	100	0	0	6	0	7,587	0.67
November	1994	7,587	100	0	0	6	0	7,430	-3.39
December	1994	7,430	100	0	169	6	0	7,581	0.69
January	1995	7,581	100	0	0	6	0	7,896	2.83
February	1995	7,896	100	0	0	6	0	8,243	3.13
March	1995	8,243	100	0	67	6	0	8,511	2.03
April	1995	8,511	100	0	0	6	0	8,805	2.28
May	1995	8,805	100	0	0	6	0	9,225	3.63
June	1995	9,225	100	0	66	6	0	9,495	1.85
July	1995	9,495	100	0	0	6	0	9,892	3.13
August	1995	9,892	100	0	0	6	0	10,035	0.43
September	1995	10,035	100	0	68	6	0	10,409	2.73
October	1995	10,409	100	0	0	6	0	10,421	-0.85
November	1995	10,421	100	0	0	6	0	10,982	4.42
December	1995	10,982	100	0	637	6	0	11,335	2.31
January	1996	11,335	100	0	0	6	0	11,645	1.85
February	1996	11,645	100	0	0	6	0	11,758	0.11
March	1996	11,758	100	0	181	6	0	12,017	1.35
April	1996	12,017	100	0	0	6	0	12,111	-0.05
May	1996	12,111	100	0	0	6	0	12,413	1.67
June	1996	12,413	100	0	66	6	0	12,517	0.03
July	1996	12,517	100	0	0	6	0	12,131	-3.88
August	1996	12,131	100	0	0	6	0	12,555	2.67
September	1996	12,555	100	0	68	6	0	13,089	3.45
October	1996	13,089	100	0	0	6	0	13,460	2.07

Hypothetical Portfolio Illustration

03-31-1991 to 03-31-2011

Investment Detail

Period		Beginning Balance	New Investment	Distribution/ Withdrawal	Total Reinvest	Charges & Fees	Taxes Due	Market Value	Total Return %
Totals		0	26,000	0	34,210	1,371	0	72,791	8.75
November	1996	13,460	100	0	0	6	0	14,391	6.17
December	1996	14,391	100	0	712	6	0	14,332	-1.10
January	1997	14,332	100	0	0	6	0	14,937	3.52
February	1997	14,937	100	0	0	6	0	15,181	0.97
March	1997	15,181	100	0	322	6	0	14,832	-2.96
April	1997	14,832	100	0	0	6	0	15,413	3.25
May	1997	15,413	100	0	0	6	0	16,399	5.74
June	1997	16,399	100	0	78	6	0	17,042	3.31
July	1997	17,042	100	0	0	6	0	18,359	7.14
August	1997	18,359	100	0	0	6	0	17,680	-4.24
September	1997	17,680	100	0	80	6	0	18,586	4.56
October	1997	18,586	100	0	0	6	0	18,219	-2.51
November	1997	18,219	100	0	0	6	0	18,662	1.88
December	1997	18,662	100	0	2,018	6	0	19,030	1.44
January	1998	19,030	100	0	0	6	0	19,230	0.52
February	1998	19,230	100	0	0	6	0	20,328	5.19
March	1998	20,328	100	0	296	6	0	21,385	4.71
April	1998	21,385	100	0	0	6	0	21,614	0.61
May	1998	21,614	100	0	0	6	0	21,654	-0.28
June	1998	21,654	100	0	95	6	0	22,092	1.56
July	1998	22,092	100	0	0	6	0	21,911	-1.27
August	1998	21,911	100	0	0	6	0	19,933	-9.48
September	1998	19,933	100	0	97	6	0	20,625	2.97
October	1998	20,625	100	0	0	6	0	21,728	4.86
November	1998	21,728	100	0	0	6	0	22,752	4.25
December	1998	22,752	100	0	1,256	6	0	23,482	2.77
January	1999	23,482	100	0	0	6	0	23,546	-0.15
February	1999	23,546	100	0	0	6	0	22,887	-3.22
March	1999	22,887	100	0	353	6	0	23,208	0.96
April	1999	23,208	100	0	0	6	0	24,537	5.30
May	1999	24,537	100	0	0	6	0	24,755	0.48
June	1999	24,755	100	0	109	6	0	25,612	3.06
July	1999	25,612	100	0	0	6	0	25,424	-1.12
August	1999	25,424	100	0	0	6	0	25,141	-1.51
September	1999	25,141	100	0	111	6	0	24,554	-2.73
October	1999	24,554	100	0	0	6	0	25,858	4.91
November	1999	25,858	100	0	0	6	0	26,368	1.59
December	1999	26,368	100	0	2,376	6	0	27,026	2.12
January	2000	27,026	100	0	0	6	0	26,767	-1.33
February	2000	26,767	100	0	0	6	0	27,039	0.64
March	2000	27,039	100	0	863	6	0	29,074	7.16
April	2000	29,074	100	0	0	6	0	29,168	-0.02
May	2000	29,168	100	0	0	6	0	29,559	1.00
June	2000	29,559	100	0	167	6	0	29,669	0.03

Hypothetical Portfolio Illustration

03-31-1991 to 03-31-2011

Investment Detail

Period		Beginning Balance	New Investment	Distribution/ Withdrawal	Total Reinvest	Charges & Fees	Taxes Due	Market Value	Total Return %
Totals		0	26,000	0	34,210	1,371	0	72,791	8.75
July	2000	29,669	100	0	0	6	0	30,100	1.12
August	2000	30,100	100	0	0	6	0	32,371	7.21
September	2000	32,371	100	0	169	6	0	32,676	0.63
October	2000	32,676	100	0	0	6	0	32,960	0.56
November	2000	32,960	100	0	0	6	0	32,332	-2.21
December	2000	32,332	100	0	2,766	6	0	33,725	4.00
January	2001	33,725	100	0	0	6	0	33,861	0.11
February	2001	33,861	100	0	0	6	0	32,866	-3.23
March	2001	32,866	100	0	884	6	0	32,059	-2.76
April	2001	32,059	100	0	0	6	0	33,325	3.64
May	2001	33,325	100	0	0	6	0	34,115	2.07
June	2001	34,115	100	0	240	6	0	33,927	-0.85
July	2001	33,927	100	0	0	6	0	34,594	1.67
August	2001	34,594	100	0	0	6	0	33,716	-2.83
September	2001	33,716	100	0	244	6	0	32,110	-5.06
October	2001	32,110	100	0	0	6	0	32,115	-0.30
November	2001	32,115	100	0	0	6	0	33,782	4.88
December	2001	33,782	100	0	415	6	0	34,116	0.69
January	2002	34,116	100	0	0	6	0	34,119	-0.28
February	2002	34,119	100	0	0	6	0	34,443	0.66
March	2002	34,443	100	0	499	6	0	35,496	2.77
April	2002	35,496	100	0	0	6	0	35,263	-0.94
May	2002	35,263	100	0	0	6	0	35,404	0.12
June	2002	35,404	100	0	197	6	0	33,811	-4.78
July	2002	33,811	100	0	0	6	0	31,914	-5.91
August	2002	31,914	100	0	0	6	0	32,341	1.03
September	2002	32,341	100	0	200	6	0	30,291	-6.65
October	2002	30,291	100	0	0	6	0	31,541	3.80
November	2002	31,541	100	0	0	6	0	32,650	3.20
December	2002	32,650	100	0	203	6	0	32,365	-1.18
January	2003	32,365	100	0	0	6	0	31,971	-1.53
February	2003	31,971	100	0	0	6	0	31,771	-0.94
March	2003	31,771	100	0	207	6	0	31,774	-0.30
April	2003	31,774	100	0	0	6	0	33,656	5.61
May	2003	33,656	100	0	0	6	0	35,742	5.90
June	2003	35,742	100	0	210	6	0	36,142	0.84
July	2003	36,142	100	0	0	6	0	36,438	0.54
August	2003	36,438	100	0	0	6	0	36,936	1.09
September	2003	36,936	100	0	213	6	0	37,038	0.01
October	2003	37,038	100	0	0	6	0	38,306	3.15
November	2003	38,306	100	0	0	6	0	38,810	1.05
December	2003	38,810	100	0	215	6	0	40,820	4.92
January	2004	40,820	100	0	0	6	0	41,483	1.38
February	2004	41,483	100	0	0	6	0	42,406	1.98

Hypothetical Portfolio Illustration

03-31-1991 to 03-31-2011

Investment Detail

Period		Beginning Balance	New Investment	Distribution/ Withdrawal	Total Reinvest	Charges & Fees	Taxes Due	Market Value	Total Return %
Totals		0	26,000	0	34,210	1,371	0	72,791	8.75
March	2004	42,406	100	0	218	6	0	41,940	-1.34
April	2004	41,940	100	0	0	6	0	41,407	-1.51
May	2004	41,407	100	0	0	6	0	41,658	0.37
June	2004	41,658	100	0	221	6	0	42,501	1.78
July	2004	42,501	100	0	0	6	0	42,066	-1.26
August	2004	42,066	100	0	0	6	0	42,320	0.37
September	2004	42,320	100	0	223	6	0	43,117	1.65
October	2004	43,117	100	0	0	6	0	43,694	1.11
November	2004	43,694	100	0	0	6	0	45,184	3.18
December	2004	45,184	100	0	378	6	0	46,851	3.47
January	2005	46,851	100	0	0	6	0	46,184	-1.64
February	2005	46,184	100	0	0	6	0	47,422	2.46
March	2005	47,422	100	0	854	6	0	46,949	-1.21
April	2005	46,949	100	0	0	6	0	46,821	-0.49
May	2005	46,821	100	0	0	6	0	47,752	1.78
June	2005	47,752	100	0	254	6	0	48,325	0.99
July	2005	48,325	100	0	0	6	0	49,715	2.67
August	2005	49,715	100	0	0	5	0	50,149	0.67
September	2005	50,149	100	0	257	5	0	50,842	1.18
October	2005	50,842	100	0	0	5	0	50,425	-1.02
November	2005	50,425	100	0	0	5	0	51,262	1.46
December	2005	51,262	100	0	2,020	5	0	51,703	0.67
January	2006	51,703	100	0	0	5	0	52,752	1.83
February	2006	52,752	100	0	0	5	0	52,847	-0.01
March	2006	52,847	100	0	1,115	5	0	53,157	0.40
April	2006	53,157	100	0	0	5	0	54,109	1.60
May	2006	54,109	100	0	0	5	0	53,408	-1.48
June	2006	53,408	100	0	304	5	0	53,447	-0.11
July	2006	53,447	100	0	0	5	0	54,470	1.73
August	2006	54,470	100	0	0	5	0	55,185	1.13
September	2006	55,185	100	0	317	5	0	56,408	2.03
October	2006	56,408	100	0	0	5	0	57,567	1.88
November	2006	57,567	100	0	0	5	0	58,351	1.19
December	2006	58,351	100	0	2,077	5	0	59,408	1.64
January	2007	59,408	100	0	0	5	0	59,960	0.76
February	2007	59,960	100	0	0	5	0	59,403	-1.10
March	2007	59,403	100	0	593	5	0	60,037	0.90
April	2007	60,037	100	0	0	5	0	62,447	3.85
May	2007	62,447	100	0	0	5	0	64,132	2.54
June	2007	64,132	100	0	365	5	0	63,398	-1.30
July	2007	63,398	100	0	0	5	0	61,623	-2.96
August	2007	61,623	100	0	0	5	0	62,254	0.86
September	2007	62,254	100	0	369	5	0	63,529	1.89
October	2007	63,529	100	0	0	5	0	64,367	1.16

Hypothetical Portfolio Illustration

03-31-1991 to 03-31-2011

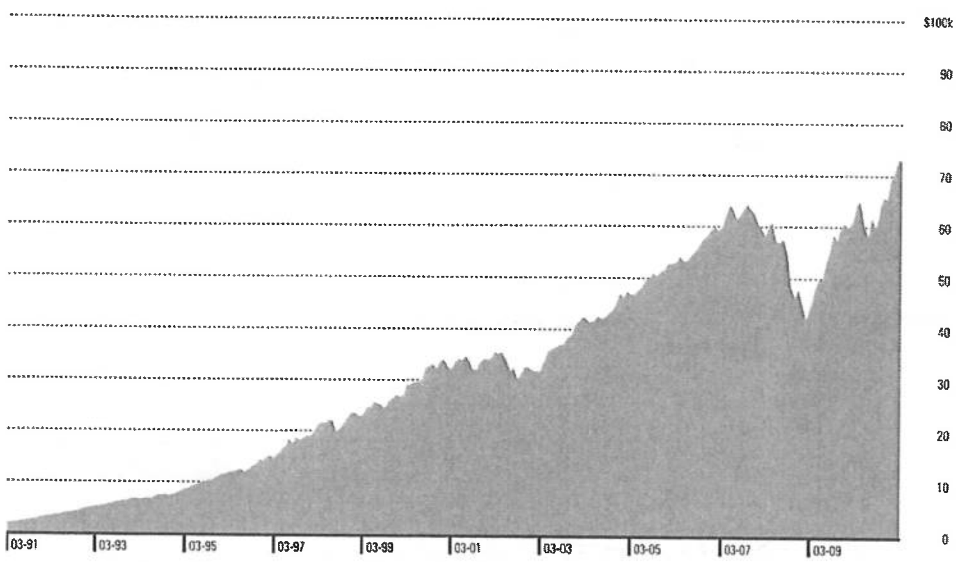
Investment Detail

Period		Beginning Balance	New Investment	Distribution/ Withdrawal	Total Reinvest	Charges & Fees	Taxes Due	Market Value	Total Return %
Totals		0	26,000	0	34,210	1,371	0	72,791	8.75
November	2007	64,367	100	0	0	5	0	63,381	-1.69
December	2007	63,381	100	0	2,558	5	0	62,488	-1.57
January	2008	62,488	100	0	0	5	0	60,533	-3.29
February	2008	60,533	100	0	0	5	0	59,496	-1.88
March	2008	59,496	100	0	463	5	0	58,354	-2.09
April	2008	58,354	100	0	0	5	0	60,454	3.43
May	2008	60,454	100	0	0	5	0	60,693	0.23
June	2008	60,693	100	0	392	5	0	57,149	-6.00
July	2008	57,149	100	0	0	5	0	57,027	-0.39
August	2008	57,027	100	0	0	5	0	57,630	0.88
September	2008	57,630	100	0	397	5	0	54,765	-5.15
October	2008	54,765	100	0	0	6	0	48,481	-11.66
November	2008	48,481	100	0	0	6	0	46,372	-4.56
December	2008	46,372	100	0	402	6	0	47,983	3.26
January	2009	47,983	100	0	0	6	0	44,953	-6.52
February	2009	44,953	100	0	0	6	0	42,215	-6.31
March	2009	42,215	100	0	315	6	0	44,738	5.74
April	2009	44,738	100	0	0	6	0	47,548	6.06
May	2009	47,548	100	0	0	6	0	49,835	4.60
June	2009	49,835	100	0	319	6	0	49,865	-0.14
July	2009	49,865	100	0	0	5	0	53,550	7.19
August	2009	53,550	100	0	0	5	0	55,864	4.13
September	2009	55,864	100	0	246	5	0	58,587	4.70
October	2009	58,587	100	0	0	5	0	57,526	-1.98
November	2009	57,526	100	0	0	5	0	60,093	4.29
December	2009	60,093	100	0	248	5	0	60,594	0.67
January	2010	60,594	100	0	0	5	0	59,989	-1.16
February	2010	59,989	100	0	0	5	0	61,331	2.07
March	2010	61,331	100	0	312	5	0	64,397	4.84
April	2010	64,397	100	0	0	5	0	65,043	0.85
May	2010	65,043	100	0	0	5	0	60,734	-6.78
June	2010	60,734	100	0	276	5	0	58,344	-4.10
July	2010	58,344	100	0	0	5	0	61,848	5.83
August	2010	61,848	100	0	0	5	0	59,880	-3.34
September	2010	59,880	100	0	280	5	0	63,598	6.04
October	2010	63,598	100	0	0	5	0	65,774	3.26
November	2010	65,774	100	0	0	5	0	65,549	-0.49
December	2010	65,549	100	0	335	5	0	69,358	5.66
January	2011	69,358	100	0	0	5	0	71,068	2.32
February	2011	71,068	100	0	0	5	0	73,266	2.95
March	2011	73,266	100	0	312	5	0	72,791	-0.78

Hypothetical Portfolio Illustration Continued

03-31-1991 to 03-31-2011

Security Summary



% of Total Market Value

Start 03-31-1991

End 03-31-2011

Investment Assumptions

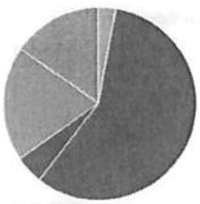
Investment Name	Holding Period Start	Holding Period End	Initial Investment Amount	Subsequent Invest/Withdwl Amount	Subsequent Invest/Withdwl Freq	Reinvest Distributions Income	Reinvest Distributions Cap Gains	Liquidate	Rebalance %	Charges and Fees Front Load	Charges and Fees Annual Fee%	Charges and Fees Deferred Load Amount%	Charges and Fees Period Years	Market Value End $
Invesco Van Kampen Equity and Income A (USD)	03-91	03-11	2,000	100	Mon	Y	Y	N	—	5.50%	0.00	0.00-0.00	—	72,791

Portfolio Snapshot

Portfolio Value
$72,791

Benchmark
S&P 500 TR (USD)

Analysis 03-31-2011



Asset Allocation	Portfolio Net %	Bmark Net %
Cash	2.94	0.00
US Stocks	57.78	99.90
Non-US Stocks	4.79	0.10
Bonds	19.24	0.00
Other/Not Clsfd	**15.25**	**0.00**

Morningstar Equity Style Box %

Total Stock Holdings 78

	Value	Core	Growth
Large	41	36	17
Mid	3	2	1
Small	0	0	0

0-10 10-25 25-50 >50

Morningstar Fixed Income Style Box %

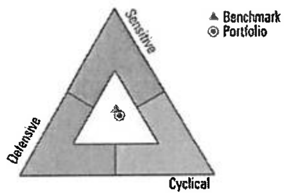
Total Bond Holdings 199

	Ltd	Mod	Ext
High	0	0	0
Med	0	100	0
Low	0	0	0

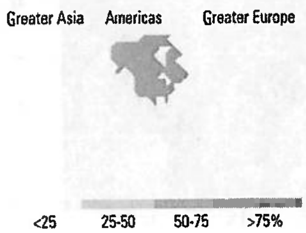
0-10 10-25 25-50 >50

Stock Analysis 03-31-2011

Stock Sectors



World Regions

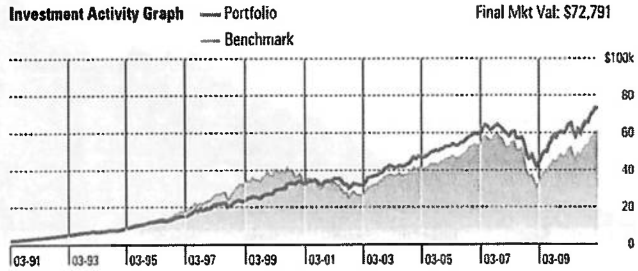


% of Stocks	Portfolio %	Bmark %
Cyclical	**36.43**	**28.50**
Basic Matls	3.28	2.96
Consumer Cycl	11.21	9.25
Financial Svs	21.94	14.67
Real Estate	0.00	1.62
Sensitive	**39.10**	**46.85**
Commun Svs	6.10	4.25
Energy	14.79	13.01
Industrials	9.87	12.93
Technology	8.34	16.66
Defensive	**24.47**	**24.65**
Consumer Def	10.71	10.73
Healthcare	9.71	10.79
Utilities	4.05	3.13
Not Classified	**0.00**	**0.00**

% of Stocks	Portfolio %	Bmark %
Greater Europe	**6.44**	**0.10**
United Kingdom	3.79	0.00
Europe-Developed	2.65	0.10
Europe-Emerging	0.00	0.00
Africa/Middle East	0.00	0.00
Americas	**92.35**	**99.91**
North America	92.35	99.91
Latin America	0.00	0.00
Greater Asia	**1.21**	**0.00**
Japan	1.21	0.00
Australasia	0.00	0.00
Asia-Developed	0.00	0.00
Asia-Emerging	0.00	0.00
Not Classified	**0.00**	**0.00**

Performance 03-31-2011

Investment Activity Graph



Trailing Returns	3 Mo	1 Yr	3 Yr	5 Yr	10 Yr
Portfolio Return	4.51	11.08	5.69	4.45	5.86
Benchmark Return	5.92	15.81	2.88	2.87	3.58
+/- Benchmark Return	-1.41	-4.73	2.81	1.58	2.28

Best/Worst Time Periods	Best %	Worst %
3 Months	17.30 (Mar 09-May 09)	-20.02 (Sep 08-Nov 08)
1 Year	41.93 (Mar 09-Feb 10)	-30.76 (Mar 08-Feb 09)
3 Years	24.30 (Apr 95-Mar 98)	-9.55 (Mar 06-Feb 09)

Portfolio Yield (03-31-2011)	Yield %
12-Month Yield	1.68

Performance Disclosure

The performance data quoted represents past performance and does not guarantee future results. The investment return and principal value of an investment will fluctuate thus an investor's shares, when redeemed, may be worth more or less than their original cost. Current performance may be lower or higher than return data quoted herein. For performance data current to the most recent month-end, please visit http://advisor.morningstar.com/familyinfo.asp.

See Disclosure Page for Standardized Returns.

Holdings 03-31-2011

Top 1 holding out of 1	Symbol	Type	Holding Value $	% Assets
Invesco Van Kampen Equity and Income A (USD)	ACEIX	MF	72,791	100.00

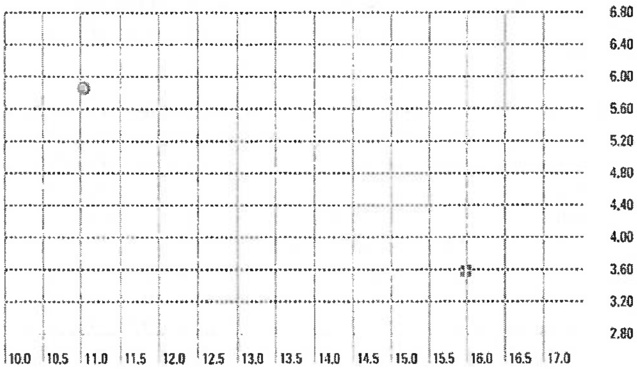
Portfolio Snapshot

Portfolio Value
$72,791

Benchmark
S&P 500 TR (USD)

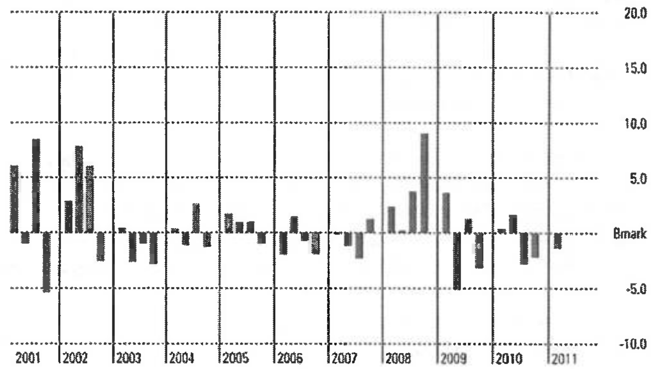
Risk Analysis 03-31-2011

Risk/Reward Scatterplot Portfolio Holdings Bmark 10 Year Mean



10 Year Standard Deviation

Performance History Graph Portfolio Quarterly returns +/- Benchmark in %



Risk and Return Statistics	3 Yr Portfolio	3 Yr Bmark	5 Yr Portfolio	5 Yr Bmark	10 Yr Portfolio	10 Yr Bmark
Standard Deviation	16.04	21.89	13.03	17.87	11.04	15.98
Mean	5.69	2.88	4.45	2.87	5.86	3.58
Sharpe Ratio	0.39	0.20	0.23	0.12	0.38	0.15

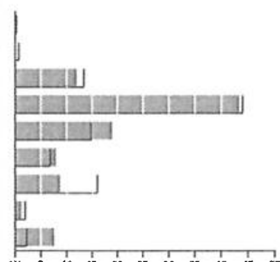
MPT Statistics	3 Yr Portfolio	5 Yr Portfolio	10 Yr Portfolio
Alpha	3.01	1.47	2.51
Beta	0.72	0.71	0.66
R-Squared	96	95	91

Fundamental Analysis 03-31-2011

Market Maturity % of Stocks	Portfolio	Bmark
Developed Markets	100.00	100.00
Emerging Markets	0.00	0.00
Not Available	**0.00**	**0.00**

Geometric Avg Capitalization ($Mil)	
Portfolio	44,879.69
Benchmark	50,179.19

Type Weightings % of Stocks	Portfolio	Bmark
High Yield	0.00	0.23
Distressed	0.00	0.67
Hard Asset	11.94	13.30
Cyclical	43.35	43.95
Slow Growth	19.06	14.82
Classic Growth	8.02	6.74
Aggressive Growth	8.80	16.09
Speculative Growth	1.23	1.98
Not Available	**7.62**	**2.22**



Valuation Multiples	Portfolio	Bmark
Price/Earnings	11.06	16.13
Price/Book	1.73	2.26
Price/Sales	1.12	1.39
Price/Cash Flow	7.68	9.50

Profitability % of Stocks	Portfolio 2011-03	Bmark 2011-03
Net Margin	11.19	12.90
ROE	15.50	20.92
ROA	5.86	8.47
Debt/Capital	38.03	35.67

Fund Statistics	
Potential Cap Gains Exposure	7.85
Avg Net Expense Ratio	0.78
Avg Gross Expense Ratio	0.78

Credit Quality Breakdown	% of Bonds
AAA	55.11
AA	10.96
A	13.73
BBB	20.20
BB	0.00
B	0.00
Below B	0.00
NR/NA	0.00

Interest Rate Risk	Portfolio
Avg Eff Maturity	7.40
Avg Eff Duration (total portfolio)	4.85
Avg Credit Quality	—
Avg Wtd Coupon	3.55

Portfolio Snapshot

Portfolio Value
$72,791

Benchmark
S&P 500 TR (USD)

Standardized and Tax Adjusted Returns

The performance data quoted represents past performance and does not guarantee future results. The investment return and principal value of an investment will fluctuate thus an investor's shares, when redeemed, may be worth more or less than their original cost. Current performance may be lower or higher than return data quoted herein. For performance data current to the most recent month-end please visit http://advisor.morningstar.com/familyinfo.asp

An investment in a money-market vehicle is not insured or guaranteed by the FDIC or any other government agency. The current yield quotation reflects the current earnings of the money market more closely than the total return quotation. Although money markets seek to preserve the value of your investment at $1.00 per share, it is possible to lose money by investing in them.

Standardized Returns assume reinvestment of dividends and capital gains. It depicts performance without adjusting for the effects of taxation, but are adjusted to reflect sales charges and ongoing fund expenses.

If adjusted for taxation, the performance quoted would be significantly reduced.

For variable annuities, additional expenses will be taken in account, including M&E risk charges, fund-level expenses such as management fees and operating fees, and contract-level administration fees, charges such as surrender, contract and sales charges.

After-tax returns are calculated using the highest individual federal marginal income tax rates, and do not reflect the impact of state and local taxes. Actual after tax returns depend on the investor's tax situation and may differ from those shown. The after tax returns shown are not relevant to investors who hold their fund shares through tax-deferred arrangements such as 401(k) plans or an IRA. After-tax returns exclude the effects of either the alternative minimum tax or phase-out of certain tax credits. Any taxes due are as of the time the distributions are made, and the taxable amount and tax character of each distribution is as specified by the fund on the dividend declaration date. Due to foreign tax credits or realized capital losses, after-tax returns may be greater than before tax returns. After-tax returns for exchange-traded funds are based on net asset value.

Annualized returns 03-31-2011

Standardized Returns (%)	7-day Yield	1Yr	5Yr	10Yr	Since Inception	Inception Date	**Max Front Load %**	**Max Back Load %**	**Net Exp Ratio %**	**Gross Exp Ratio %**
Invesco Van Kampen Equity and Income A (USD)	—	4.96	3.23	5.31	10.30	08-03-1960	**5.50**	**NA**	**0.78**	**0.78**
BarCap US Agg Bond TR USD	—	**5.12**	**6.03**	**5.56**	—	—				
MSCI EAFE NR USD	—	**10.42**	**1.30**	**5.39**	—	—				
S&P 500 TR	—	**15.65**	**2.62**	**3.29**	—	—				
USTREAS T-Bill Auction Ave 3 Mon	—	**0.15**	**2.08**	**2.14**	—	—				

Return after Tax (%)	On Distribution					On Distribution and Sales of Shares			
	1Yr	5Yr	10Yr	Since Inception	Inception Date	1Yr	5Yr	10Yr	Since Inception
Invesco Van Kampen Equity and Income A (USD)	4.28	2.14	4.18	6.37	08-03-1960	3.18	2.20	3.99	6.27

Portfolio Snapshot

Portfolio Value
$72,791

Benchmark
S&P 500 TR (USD)

Illustration Returns

Total 1 holding as of 03-31-2011	Symbol	Type	Holdings Date	% of Assets	Holding Value $	7-day Yield	1 Yr Ret %	3 Yr Ret %	5 Yr Ret %	10 Yr Ret %
Invesco Van Kampen Equity and Income A (USD)	ACEIX	MF	12-2010	100.00	72,791	—	11.08	5.69	4.45	5.86

Performance Disclosure

The performance data quoted represents past performance and does not guarantee future results. The investment return and principal value of an investment will fluctuate thus an investor's shares, when redeemed, may be worth more or less than their original cost. Current performance may be lower or higher than return data quoted herein. For performance data current to the most recent month-end, please visit http://advisor.morningstar.com/familyinfo.asp.

See Disclosure Page for Standardized Returns.

PFS INVESTMENTS INC.

Hypothetical Report Disclosure Statement

General

This is an illustration of a simulated investment and assumes the portfolio holding(s) were purchased on the first day of the period indicated. Sales and tax charges, including those required in the event of transfers between assets, are taken into account at the rates shown and may be higher or lower than what an investor would have actually paid had the investments been purchased then or now. The performance data represents past performance and is not indicative of future results. Principal value and investment returns will fluctuate, and an investor's shares/units when redeemed may be worth more or less than the original investment.

The underlying holdings of the portfolio are not federally or FDIC-insured and are not deposits or obligations of, or guaranteed by, any financial institution. Investment in securities involve investment risks including possible loss of principal and fluctuation in value.

The investment returns do not reflect active trading and do not necessarily reflect the results that might have been achieved by active management of the account. The investment returns of other clients of the adviser may differ materially from the investment portrayed.

The information contained in this report is from the most recent information available to Morningstar as of the release date, and may or may not be an accurate reflection of the current composition of the securities included in the portfolio. There is no assurance that the weightings, composition and ratios will remain the same.

Pre-inception Returns

The analysis in this report may be based, in part, on adjusted historical returns for periods prior to the fund's actual inception. These calculated returns reflect the historical performance of the oldest share class of the fund, adjusted to reflect the fees and expenses of this share class. These fees and expenses are referenced in the report's Charges and Fees section.

When pre-inception data are presented in the report, the header at the top of the report will indicate this.

While the inclusion of pre-inception data provides valuable insight into the probable long-term behavior of newer share classes of a fund, investors should be aware that an adjusted historical return can only provide an approximation of that behavior. For example, the fee structures between a retail share class will vary from that of an institutional share class, as retail shares tend to have higher operating expenses and sales charges. These adjusted historical returns are not actual returns. Calculation methodologies utilized by Morningstar may differ from those applied by other entities, including the fund itself.

The investment returns do not necessarily reflect the deduction of all investment advisory fees. Client investment returns may be reduced if additional fees are incurred.

Performance for closed-end and exchange-traded funds is calculated based on the fund's end of the day market prices as reported by the New York Stock Exchange. Separate account performance is based on the mean experience of an investor in the account.

This illustration may reflect the results of systematic investments and/or withdrawals. Systematic investment does not ensure a profit, nor does it protect the investor against a loss in a declining market. Also, systematic investing will not keep an investor from losing money if shares are sold when the market is down.

Investment Summary Graph

The investment summary graph plots the approximate market value of the security or portfolio over the investing horizon. It may also include the total investment assumed in the illustration and/or a benchmark. Total investment includes dollar inflows and outflows, including inflows representing noted taxes and annual fees paid out of pocket. If a benchmark index is included on a graph, it assumes a similar pattern of investment/withdrawal as for the security or portfolio. Taxes and transaction costs are aslo applied to the benchmark index. Note that direct investment in an index is not possible. Indexes are unmanaged portfolios representing different asset classes, with varying levels of associated risk. The benchmark index included in the graph may or may not represent an appropriate or accurate comparison with the security or portfolio illustrated.

Standardized Returns

For ETFs, the standardized returns reflect performance, both at market price and NAV price, without adjusting for the effects of taxation or brokers commissions. These returns are adjusted to reflect all ongoing ETF expenses and assume reinvestment of dividends and capital gains. If adjusted, the effects of taxation would reduce the performance quoted.

For HOLDs, the standardized returns reflect performance at market price, without adjusting for the effects of taxation or brokers commissions. These returns are adjusted to reflect all ongoing expenses and assume reinvestment of dividends and capital gains. If adjusted, the effects of taxation would reduce the performaance quoted.

For money market mutual funds, standardized return is total return adjusted for sales charges and reflects all ongoing fund expenses. Current 7-day yield more closely reflects the current earnings of the money market fund than the total return quotation.

For mutual funds, standardized return is total return adjusted for sales charges and reflects all ongoing fund expenses. Following this disclosure statement, standardized returns for each portfolio holding are shown.

For VA subaccounts, standardized return is total return based on its inception date within the separate account and is adjusted to reflect recurring and non-recurring charges such as surrender fees, contract charges, maximum front-end load, maximum deferred load, maximum M&E risk charge, administration fees, and actual ongoing fund-level expenses.

For VL subaccounts, standardized return is total return based on its inception date within the separate account and is adjusted to reflect recurring and non-recurring charges such as surrender fees, contract charges, maximum front-end load, maximum deferred load, maximum M&E risk charge, administration fees, and actual ongoing fund-level expenses. For VLs, additional fees specific to a VL policy such as transfer fees and cost of insurance fees, which are based on specific characteristics on an individual, are not included. If VL fees were included in the return calculations, the performance would have been significantly lower. An investor should contact their financial advisor and ask for a personalized performance illustration, either hypothetical or historical, which reflects all applicable fees and charges including the cost of insurance. Please review the prospectus and SAI for more detailed information.

Definitions of Report Terms

*Annual Fee Paid:*Your advisor was able to specify whether annual fees, if any, should be assumed paid out of pocket or from selling shares of securities held in the illustration.

*Average Annualized Return:*Average annualized money-weighted return (internal rate of return). In illustrations with time periods less than one year, this figure is not annualized.

*Capital Gains (Individual Report):*Percentage of the total market value of the holding that is attributable to the reinvestment of capital gains distributions.

*Charges & Fees (Investment Detail):*The sum of fees charged to the investor during the period, including front or deferred loads, VA charges, and annual fees.

*Cumulative Return:*The total money-weighted return of the investment over the entire time period of the illustration.

*Distribution/Withdrl:*The sum of distributions not reinvested, plus any cash withdrawals during the period.

*Income (Individual Report):*The percentage of the total market value of the holding that is attributable to the reinvestment of income or dividend distributions.

*Liquidate:*Indicates whether the advisor chose that the holding be liquidated on the end date.

*Median (Comparison Report):*The total money-weighted return (internal rate of return) of the median security in the illustration for the calendar year indicated.

*New investment:*Any new cash invested during the period.

*Principal (Individual Reports):*The percentage of the total market value of the holding that is attributable to new investment.

*Rebalance(Planning Assumptions):*Indicates whether rebalancing is used, and its frequency. "No" indicates no rebalancing. Options for rebalancing frequency are monthly, quarterly, semi-annually, and annually.

*Rebalance(Investment Assumptions):*Percentage of total asset allocation to be maintained in this holding through rebalancing.

*Securities Returns(Comparison Report):*The total money-weighted return (internal rate of return) for the holding in the calendar year indicated, taking into account cash flows, charges, and fees.

*Subsequent Invest/Withdwl:*The amount, type, and frequency of subsequent investments or withdrawals from the holding. Withdrawals are represented by a negative number. Systematic investments and withdrawals may be made monthly, quarterly, semi-annually, or annually. If "Custom," a custom schedule of investments or withdrawals was used.

*Taxes Due:*The total amount of taxes due from the investor, determined by applying specified tax rates to distributions and sale of shares during each calendar year.

*Taxes Paid:*Your advisor was able to specify whether taxes, if any, should be assumed paid out of pocket or from selling shares of securities held in the illustration.

*Net Dollars Invested:*The total out-of-pocket expense for the investor. Includes new investment, annual fees paid to advisor, and taxes due. This figure is net of withdrawals, including liquidation.

*Total Reinvest:*The sum of distributions reinvested during the period.

*Total Return %:*The total money-weighted return (internal rate of return) on investments for the period.

Portfolio Snapshot Report Disclosure Statement

General

Investment portfolios illustrated in this report can be scheduled or unscheduled. With an unscheduled portfolio, the user inputs only the portfolio holdings and their current allocations. Morningstar calculates returns using the given allocations assuming monthly rebalancing. Taxes, loads, and sales charges are not taken into account.

With "scheduled" portfolios, users input the date and amount for all investments into and withdrawals from each holding, as well as tax rates, loads, and other factors that would have affected portfolio performance. A hypothetical illustration is one type of scheduled portfolio.

Both scheduled and unscheduled portfolios are theoretical, for illustrative purposes only, and are not reflective of an investors actual experience. For both scheduled and unscheduled portfolios, the performance data given represents past performance and should not be considered indicative of future results. Principal value and investment return of stocks, mutual funds, and variable annuity/life products will fluctuate, and an investor's shares/units when redeemed will be worth more or less than the original investment. Stocks, mutual funds, and variable annuity/life products are not FDIC-insured, may lose value, and are not guaranteed by a bank or other financial institution. Portfolio statistics change over time.

Used as supplemental sales literature, the Portfolio Snapshot report must be preceded or accompanied by the fund/policy's current prospectus or equivalent. In all cases, this disclosure statement should accompany the Portfolio Snapshot report. Morningstar is not itself a FINRA-member firm.

The underlying holdings of the portfolio are not federally or FDIC-insured and are not deposits or obligations of, or guaranteed by any financial institution. Investment in securities involve investment risks including possible loss of principal and fluctuation in value.

The information contained in this report is from the most recent information available to Morningstar as of the release date, and may or may not be an accurate reflection of the current composition of the securities included in the portfolio. There is no assurance that the weightings, composition and ratios will remain the same.

Items to Note Regarding Certain Underlying Securities

A closed-end fund is an investment company, which typically makes one public offering of a fixed number of shares. Thereafter, shares are traded on a secondary market such as the New York Stock Exchange. As a result, the secondary market price may be higher or lower than the closed-end fund's net asset value (NAV). If these shares trade at a price above their NAV, they are said to be trading at a premium. Conversely, if they are trading at a price below their NAV, they are said to be trading at a discount.

An exchange-traded fund (ETF) is an investment company that typically has an investment objective of striving to achieve a similar return as a particular market index. The ETF will invest in either all or a representative sample of the securities included in the index it is seeking to imitate. Like closed-end funds, ETFs can be traded on a secondary market and thus have a market price that may be higher or lower than its net asset value. If these shares trade at a price above their NAV, they are said to be trading at a premium. Conversely, if they are trading at a price below their NAV, they are said to be trading at a discount.

A money market fund is an investment company that invests in commercial paper, banker's acceptances, repurchase agreements, government securities, certificates of deposit and other highly liquid securities, and pays money market rates of interest. Money markets are not FDIC-insured, may lose money, and are not guaranteed by a bank or other financial institution. Although the money market seeks to preserve a stable per share value (i.e. $1.00 per share), it is possible to lose money by investment in the fund.

Unit investment trust (UIT) is an investment company organized under a trust agreement between a sponsor and trustee. UITs typically purchase a fixed portfolio of securities and then sell units in the trust to investors. The major difference between a UIT and a mutual fund is that a mutual fund is actively managed, while a UIT is not. On a periodic basis, UITs usually distribute to the unit holder their pro rata share of the trust's net investment income and net realized capital gains, if any. If the trust is one that invests only in tax-free securities, then the income from the trust is also tax-free. UITs generally make one public offering of a fixed number of units. However, in some cases, the sponsor will maintain a secondary market that allows existing unit holders to sell their units and for new investors to buy units.

Variable annuities are tax-deferred investments structured to convert a sum of money into a series of payments over time. Variable annuity policies have limitations and are not viewed as short-term liquid investments. An insurance company's fulfillment of a commitment to pay a minimum death benefit, a schedule of payments, a fixed investment account guaranteed by the insurance company, or another form of guarantee depends on the claims-paying ability of the issuing insurance company. Any such guarantee does not affect or apply to the investment return or principal value of the separate account and its subaccount. The financial ratings quoted for an insurance company do not apply to the separate account and its subaccount. If the variable annuity subaccount is invested in a money-market fund, although it seeks to preserve a stable per share value (i.e. $1.00 per share), it is possible to lose money by investment in the fund.

Variable life insurance is a cash-value life insurance that has a variable cashvalue and/or death benefit depending on the investment performance of the subaccount into which premium payments are invested. Unlike traditional life insurance, variable life insurance has inherent risks associated with it, including market volatility, and is not viewed as a short-term liquid investment. For more information on a variable life product, including each subaccount, please read the current prospectus. Please note, the financial ratings noted on the report are quoted for an insurance company and do not apply to the separate account and its subaccount. If the variable life subaccount is invested in a money-market fund, although it seeks to preserve a stable per share value (i.e. $1.00 per share), it is possible to lose money by investment in the fund.

Pre-inception Returns

The analysis in this report may be based, in part, on adjusted historical returns for periods prior to the fund's actual inception. These calculated returns reflect the historical performance of the oldest share class of the fund, adjusted to reflect the fees and expenses of this share class. These fees and expenses are referenced in the report's list of holdings and again on the standardized returns page. When pre-inception data is presented in the report, the header at the top of the report will indicate this and the affected data elements will be displayed in italics.

While the inclusion of pre-inception data provides valuable insight into the probable long-term behavior of newer share classes of a fund, investors should be aware that an adjusted historical return can only provide an approximation of that behavior. For example, the fee structures between a retail share class will vary from that of an institutional share class, as retail shares tend to have higher operating expenses and sales charges. These adjusted historical returns are not actual returns. Calculation methodologies utilized by Morningstar may differ from those applied by other entities, including the fund itself.

Scheduled Portfolio Trailing Returns

Scheduled Portfolios are customized by the user to account for loads, taxes, cash flows and specific investment dates. Scheduled portfolios use the portfolio's investment history to calculate final market values and returns. For scheduled portfolios, both individual holdings and portfolio returns are internal-rate-of-return calculations that reflect the timing and dollar size of all purchases and sales. For stocks and mutual funds, sales charges and tax rates are taken into account as specified by the user (except in the pre-tax returns, which reflect the impact of sales charges but not taxes). Note that in some scheduled portfolio illustrations, dividends and capital gains distributions, if applicable, are reinvested at the end of the month in which they are made at the month-end closing price. This can cause discrepancies between calculated returns and actual investor experience.

Scheduled Portfolio Returns-Based Performance Data

For scheduled portfolios, the monthly returns used to calculate alphas, betas, R-squareds, standard deviations, Sharpe ratios and best/worst time-period data are internal rates of return.

Important VA Disclosure for Scheduled Portfolios

For variable annuity products, policy level charges (other than front-end loads, if input by the advisor) are not factored into returns. When withdrawals and liquidations are made, increases in value over the purchase price are taxed at the capital gains rate that currently is in effect. This is not reflective of the actual tax treatment for these products, which requires the entire withdrawal to be taxed at the income tax rate. If adjusted for sales charges and the effects of taxation, the subaccount returns would be reduced.

Scheduled Portfolio Investment Activity Graph

The historic portfolio values that are graphed are those used to track theportfolio when calculating returns.

Unscheduled Portfolio Returns

Monthly total returns for unscheduled portfolios are calculated by applying theending period holding weightings supplied by the user to an individual holding's monthly returns. When monthly returns are unavailable for a holding (ie. Due to it not being in existence during the historical period being reported), the remaining portfolio holdings are re-weighted to maintain consistent proportions. Inception dates are listed in the Disclosure for Standardized and Tax Adjusted Returns. Trailing returns are calculated by geometrically linking these weighted-average monthly returns. Unscheduled portfolio returns thus assume monthly rebalancing. Returns for individual holdings are simple time-weighted trailing returns. Neither portfolio returns nor holding returns are adjusted for loads or taxes, and if adjusted for, would reduce the returns stated. The returns stated assume the reinvestment of dividends and capital gains. Mutual fund returns include all ongoing fund expenses. VA/VL returns reflect subaccount level fund expenses, including M&E expenses, administration fees, and actual ongoing fund level expenses.

Unscheduled Portfolio Investment Activity Graph

The historic performance data graphed is extrapolated from the ending portfolio value based on monthly returns.

Benchmark Returns
Benchmark returns may or may not be adjusted to reflect ongoing expenses such as sales charges. An investment's portfolio may differ significantly from the securities in the benchmark.

Returns for custom benchmarks are calculated by applying user-supplied weightings to each benchmark's returns every month. Trailing returns are calculated by geometrically linking these weighted-average monthly returns. Custom benchmark returns thus assume monthly rebalancing.

Standardized Returns
For mutual funds, standardized return is total return adjusted for sales charges, and reflects all ongoing fund expenses. Following this disclosure statement, standardized returns for each portfolio holding are shown.

For money market mutual funds, standardized return is total return adjusted for sales charges and reflects all ongoing fund expenses. Current 7-day yield more closely reflects the current earnings of the money market fund than the total return quotation.

For VA subaccounts, standardized return is total return based on its inception date within the separate account and is adjusted to reflect recurring and non-recurring charges such as surrender fees, contract charges, maximum front-end load, maximum deferred load, maximum M&E risk charge, administration fees and actual ongoing fund-level expenses.

For ETFs, the standardized returns reflect performance, both at market priceand NAV price, without adjusting for the effects of taxation or brokers commissions. These returns are adjusted to reflect all ongoing ETF expenses and assume reinvestment of dividends and capital gains. If adjusted, the effects of taxation would reduce the performance quoted.

The charges and expenses used in the standardized returns are obtained from the most recent prospectus and/or shareholder report available to Morningstar. For mutual funds and VAs, all dividends and capital gains are assumed to be reinvested. For stocks, stock acquired via divestitures is assumed to be liquidated and reinvested in the original holding.

Non-Standardized Returns
For mutual funds, total return is not adjusted for sales charges and reflects all ongoing fund expenses for various time periods. These returns assume reinvestment of dividends and capital gains. If adjusted for sales charges and the effects of taxation, the mutual fund returns would be reduced. Please note these returns can include pre-inception data and if included, this data will be represented in italics.

For money market funds, total return is not adjusted for sales charges and reflects all ongoing fund expenses for various time periods. These returns assume reinvestment of dividends and capital gains. If adjusted for sales charges and the effects of taxation, the money market returns would be reduced.

For VA and VL subaccounts, non-standardized returns illustrate performance that is adjusted to reflect recurring and non-recurring charges such as surrender fees, contract charges, maximum front-end load, maximum deferred load, maximum M&E risk charge, administrative fees and underlying fund-level expenses for various time periods. Non-Standardized performance returns assume reinvestment of dividends and capital gains. If adjusted for the effects of taxation, the subaccount returns would be significantly reduced. Please note these returns can include pre-inception data and if included, this data will be represented in italics.

Investment Advisory Fees
The investment(s) returns do not necessarily reflect the deduction of all investment advisory fees. Client investment returns will be reduced if additional advisory fees are incurred such as deferred loads, redemption fees, wrap fees, or other account charges.

Investment Style
The Morningstar Style Box reveals a fund's investment style as of the date noted on this report.

For equity funds the vertical axis shows the market capitalization of the long stocks owned and the horizontal axis shows investment style (value, blend, or growth).

For fixed-income funds, the vertical axis shows the credit quality of the long bonds owned and the horizontal axis shows interest rate sensitivity as measured by a bond's effective duration.

Morningstar seeks credit rating information from fund companies on a periodic basis (e.g., quarterly). In compiling credit rating information, Morningstar instructs fund companies to only use ratings that have been assigned by a Nationally Recognized Statistical Rating Organization (NRSRO). If two NRSROs have rated a security, fund companies are to report the lowest rating; if three or more NRSROs have rated the same security differently, fund companies are to report the rating that is in the middle. For example, if NRSRO X rates a security AA-, NRSRO Y rates the same security an A and NRSRO Z rates it a BBB+, the fund company should use the credit rating of 'A' in its reporting to Morningstar. PLEASE NOTE: Morningstar, Inc. is not itself an NRSRO nor does it issue a credit rating on the fund. An NRSRO rating on a fixed-income security can change from time-to-time.

For credit quality, Morningstar combines the credit rating information provided by the fund companies with an average default rate calculation to come up with a weighted-average credit quality. The weighted-average credit quality is currently a letter that roughly corresponds to the scale used by a leading NRSRO. Bond funds are assigned a style box placement of "low", "medium", or "high" based on their average credit quality. Funds with a low credit quality are those whose weighted-average credit quality is determined to be less than "BBB-"; medium are those less than "AA-", but greater or equal to "BBB-"; and high are those with a weighted-average credit quality of "AA-" or higher. When classifying a bond portfolio, Morningstar first maps the NRSRO credit ratings of the underlying holdings to their respective default rates (as determined by Morningstar's analysis of actual historical default rates). Morningstar then averages these default rates to determine the average default rate for the entire bond fund. Finally, Morningstar maps this average default rate to its corresponding credit rating along a convex curve.

For interest-rate sensitivity, Morningstar obtains from fund companies the average effective duration. Generally, Morningstar classifies a fixed-income fund's interest-rate sensitivity based on the effective duration of the Morningstar Core Bond Index (MCBI), which is currently three years. The classification of Limited will be assigned to those funds whose average effective duration is between 25% to 75% of MCBI's average effective duration; funds whose average effective duration is between 75% to 125% of the MCBI will be classified as Moderate; and those that are at 125% or greater of the average effective duration of the MCBI will be classified as Extensive.

For municipal bond funds, Morningstar also obtains from fund companies the average effective duration. In these cases static breakpoints are utilized. These breakpoints are as follows: (i) Limited: 4.5 years or less; (ii) Moderate: more than 4.5 years but less than 7 years; and (iii) Extensive: more than 7 years. In addition, for non-US taxable and non-US domiciled fixed income funds static duration breakpoints are used: (i) Limited: less than or equal to 3.5 years; (ii) Moderate: greater than 3.5 and less than equal to 6 years; (iii) Extensive: greater than 6 years.

Risk and Return

Standard deviation is a statistical measure of the volatility of a portfolio's returns around its mean.

Mean represents the annualized geometric return for the period shown.

Sharpe ratio uses a portfolio's standard deviation and total return to determine reward per unit of risk.

Alpha measures the difference between a portfolio's actual returns and its expected performance, given its beta and the actual returns of the benchmark index. Alpha is often seen as a measurement of the value added or subtracted by a portfolio's manager.

Beta is a measure of the degree of change in value one can expect in a portfolio given a change in value in a benchmark index. A portfolio with a beta greater than one is generally more volatile than its benchmark index, and a portfolio with a beta of less than one is generally less volatile than its benchmark index.

R-squared reflects the percentage of a portfolio's movements that is explained by movements in its benchmark index, showing the degree of correlation between the portfolio and a benchmark. This figure is also helpful in assessing how likely it is that alpha and beta are statistically significant.

Fundamental Analysis

The below referenced data elements are a weighted average of the equity holdings in the portfolio.

The median market capitalization of a subaccount's equity portfolio gives you a measure of the size of the companies in which the subaccount invests.

The Price/Cash Flow ratio is a weighted average of the price/cash-flow ratios of the stocks in a subaccounts portfolio. Price/cash-flow shows the ability of a business to generate cash and acts as a gauge of liquidity and solvency.

The Price/Book ratio is a weighted average of the price/book ratios of all the stocks in the underlying fund's portfolio. The P/B ratio of a company is calculated by dividing the market price of its stock by the company's per-share book value. Stocks with negative book values are excluded from this calculation.

The Price/Earnings ratio is calculated by dividing the market value of the equity assets by the trailing 12 month earnings. The 12 month earnings value comes from multiplying the number of shares and the adjusted trailing 12 months' earnings per share for each equity asset and summing the results.

The Price/Sales ratio is a weighted average of the price/sales ratios of the stocks in the underlying fund's portfolio. The P/S ratio of a stock is calculated by dividing the current price of the stock by its trailing 12 months' revenues per share. In computing the average, Morningstar weights each portfolio holding by the percentage of equity assets it represents.

The return on assets (ROA) is the percentage a company earns on its assets in a given year. The calculation is net income divided by end-of-year total assets, multiplied by 100.

The Return on Equity (ROE) is the percentage a company earns on its shareholders' equity in a given year. The calculation is net income divided by end-of-year net worth, multiplied by 100.

Market Maturity shows the percentage of a holding's common stocks that are domiciled in developed and emerging markets.

The data elements listed below are a weighted average of the fixed income holdings in the portfolio.

Average maturity is used for holdings in the taxable fixed-income category. This is a weighted average of all the maturities of the bonds in a portfolio, computed by weighting each maturity date by the market value of the security.

Credit quality breakdowns are shown for corporate-bond holdings and depict the quality of bonds in the underlying portfolio. The report shows the percentage of fixed-income securities that fall within each credit quality rating as assigned by an NRSRO. Bonds not rated by an NRSRO are included in the not rated (NR) category.

Debt as a percentage of capital is calculated by dividing long-term debt by total capitalization (the sum of common equity plus preferred equity plus long-term debt). This figure is not provided for financial companies.

Duration is a time measure of a bonds interest-rate sensitivity.

Net Margin is a measure of profitability. It is equal to annual net income divided by revenues from the same period for the past five fiscal years, multiplied by 100.

Type Weightings divide the stocks in a given holding's portfolio into eight type designations each of which defines a broad category of investment characteristics. Not all stocks in a given holding's portfolio are assigned a type. These stocks are grouped under NA.

The data elements listed below are a weighted average of the total holdings in the portfolio.

The average expense ratio is the percentage of assets deducted each year for operating expenses, management fees, and all other asset-based costs incurred by the fund, excluding brokerage fees. Please note for mutual funds, variable annuities/life, ETF and closed-end funds we use the gross prospectus ratio as provided in the prospectus. Separate accounts and stocks are excluded from the average expense ratio.

Potential capital gains exposure is the percentage of a holdings total assets that represent capital appreciation.

Investment Risks

International/Emerging Market Equities: Investing in international securities involve special additional risks. These risks include, but are not limited to, currency risk, political risk, and risk associated with varying accounting standards. Investing in emerging markets may accentuate these risks.

Sector Strategies: Portfolios that invest exclusively in one sector or industry involve additional risks. The lack of industry diversification subjects the investor to increased industry-specific risks.

Non-Diversified Strategies: Portfolios that invest a significant percentage of assets in a single issuer involve additional risks, including share price fluctuations, because of the increased concentration of investments.

Small Cap Equities: Portfolios that invest in stocks of small companies involve additional risks. Smaller companies typically have a higher risk of failure, and are not as well established as larger blue-chip companies. Historically, smaller-company stocks have experienced a greater degree of market volatility that the overall market average.

Mid Cap Equities: Portfolios that invest in companies with market capitalization below $10 billion involve additional risks. The securities of these companies may be more volatile and less liquid than the securities of larger companies.

High-Yield Bonds: Portfolios that invest in lower-rated debt securities (commonly referred as junk bonds) involve additional risks because of the lower credit quality of the securities in the portfolio. The investor should be aware of the possible higher level of volatility, and increased risk of default.

Tax-Free Municipal Bonds: The investor should note that the income from tax-free municipal bond funds may be subject to state and local taxation and the Alternative Minimum Tax.

Bonds: Bonds are subject to interest rate risk. As the prevailing level of bond interest rates rise, the value of bonds already held in a portfolio decline. Portfolios that hold bonds are subject to declines and increases in value due to general changes in interest rates.

HOLDRs: The investor should note that these are narrow industry-focused products that, if the industry is hit by hard times, will lack diversification and possible loss of investment would be likely. These securities can trade at a discount to market price, ownership is of a fractional share interest, the underlying investments may not be representative of the particular industry, the HOLDR might be delisted from the AMEX if the number of underlying companies drops below nine, and the investor may experience trading halts.

Hedge Funds: The investor should note that hedge fund investing involves specialized risks that are dependent upon the type of strategies undertaken by the manager. This can include distressed or event-driven strategies, long/short strategies, using arbitrage (exploiting price inefficiencies), international investing, and use of leverage, options and/or derivatives. Although the goal of hedge fund managers may be to reduce volatility and produce positive absolute return under a variety of market conditions, hedge funds may involve a high degree of risk and are suitable only for investors of substantial financial means who could bear the entire loss of their investment.

Bank Loan/Senior Debt: Bank loans and senior loans are impacted by the risks associated with fixed income in general, including interest rate risk and default risk. They are often non-investment grade; therefore, the risk of default is high. These securities are also relatively illiquid. Managed products that invest in bank loans/senior debt are often highly leveraged, producing a high risk of return volatility.

Short Positions: When a short position moves in an unfavorable way, the losses are theoretically unlimited. The broker may demand more collateral and a manager might have to close out a short position at an inopportune time to limit further losses.

Long-Short: Due to the strategies used by long-short funds, which may include but are not limited to leverage, short selling, short-term trading, and investing in derivatives, these funds may have greater risk, volatility, and expenses than those focusing on traditional investment strategies.

Liquidity Risk: Closed-end fund, ETF, and HOLDR trading may be halted due to market conditions, impacting an investor's ability to sell a fund.

Market Price Risk: The market price of ETFs, HOLDRs, and closed-end funds traded on the secondary market is subject to the forces of supply and demand and thus independent of the NAV. This can result in the market price trading at a premium or discount to the NAV which will affect an investor's value.

Market Risk: The market prices of ETF's and HOLDRs can fluctuate as a result of several factors, such as security-specific factors or general investor sentiment. Therefore, investors should be aware of the prospect of market fluctuations and the impact it may have on the market price.

Target-Date Funds: Target-date funds typically invest in other mutual funds and are designed for investors who are planning to retire during the target date year. The fund's target date is the approximate date of when investors expect to begin withdrawing their money. Target-date fund's investment objective/strategy typically becomes more conservative over time primarily by reducing its allocation to equity mutual funds and increasing its allocations in fixed-income mutual funds. An investor's principal value in a target-date fund is not guaranteed at anytime, including at the fund's target date.

High double- and triple-digit returns were the result of extremely favorable market conditions, which may not continue to be the case. High returns for short time periods must not be a major factor when making investment decisions.

Benchmark Disclosure

BarCap US Agg Bond TR USD

This index is composed of the BarCap Government/Credit Index, the Mortgage-Backed Securities Index, and the Asset-Backed Securities Index. The returns we publish for the index are total returns, which include reinvestment of dividends.

MSCI EAFE NR USD

This Europe, Australasia, and Far East index is a market-capitalization-weighted index of 21 non-U.S., industrialized country indexes.

S&P 500 TR

A market capitalization-weighted index of 500 widely held stocks often used as a proxy for the stock market. TR (Total Return) indexes include daily reinvestment of dividends.

USTREAS T-Bill Auction Ave 3 Mon

Three-month T-bills are government-backed short-term investments considered to be risk-free and as good as cash because the maturity is only three months. Morningstar collects yields on the T-bill on a weekly basis from the Wall Street Journal.

PFS INVESTMENTS INC.

Member of FINRA

Hypo Report

April 10, 2011

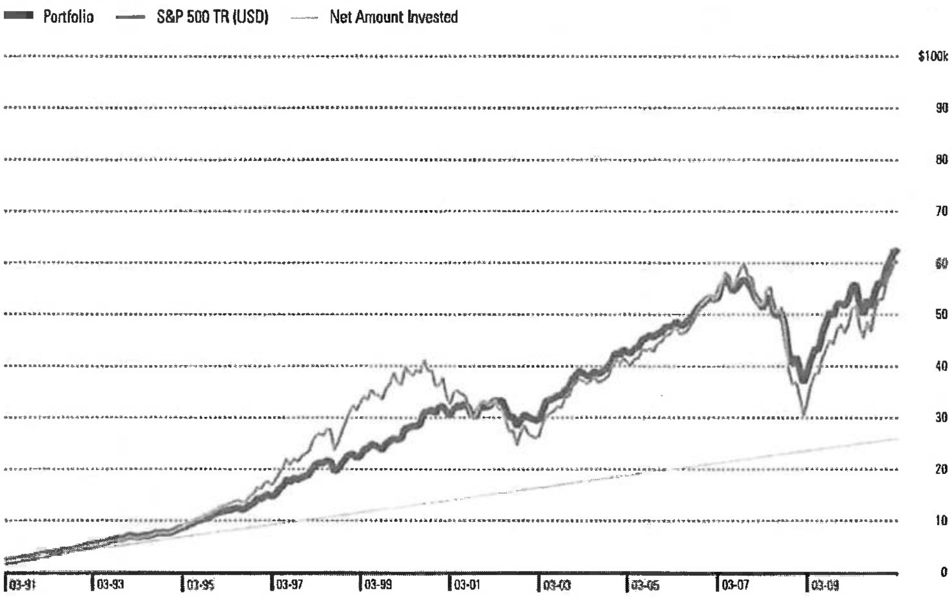
Prepared Especially For:

Your Representative:

Hypothetical Portfolio Illustration

03-31-1991 to 03-31-2011

Portfolio Summary



Planning Assumptions

Currency	USD
Rebalance	None
Annual Fee Paid	Sale of Shares
Federal Income Tax Rate	0%
Capital Gain Tax Rate	0%
State Tax Rate	0%
Tax Paid	Out of Pocket

Performance

Net Amount Invested	$26,000
Final Market Value	$61,860
Average Annualized Return	7.45%
Cumulative Return	321.14%

Investment Detail

Period		Beginning Balance	New Investment	Distribution/ Withdrawal	Total Reinvest	Charges & Fees	Taxes Due	Market Value	Total Return %
Totals		0	26,000	0	31,887	8,839	0	61,860	7.45
March	1991	0	2,000	0	0	8	0	1,993	-0.37
April	1991	1,993	100	0	0	0	0	2,097	0.22
May	1991	2,097	100	0	0	0	0	2,264	3.20
June	1991	2,264	100	0	25	9	0	2,284	-3.51
July	1991	2,284	100	0	0	0	0	2,479	4.16
August	1991	2,479	100	0	0	0	0	2,645	2.66
September	1991	2,645	100	0	29	10	0	2,735	-0.38
October	1991	2,735	100	0	0	0	0	2,883	1.74
November	1991	2,883	100	0	0	0	0	2,915	-2.37
December	1991	2,915	100	0	42	12	0	3,233	7.49
January	1992	3,233	100	0	0	0	0	3,313	-0.63
February	1992	3,313	100	0	0	0	0	3,468	1.66
March	1992	3,468	100	0	36	13	0	3,498	-2.02
April	1992	3,498	100	0	0	0	0	3,664	1.89
May	1992	3,664	100	0	0	0	0	3,816	1.44
June	1992	3,816	100	0	39	15	0	3,863	-1.40
July	1992	3,863	100	0	0	0	0	4,091	3.32
August	1992	4,091	100	0	0	0	0	4,158	-0.81
September	1992	4,158	100	0	42	16	0	4,292	0.83
October	1992	4,292	100	0	0	0	0	4,410	0.40
November	1992	4,410	100	0	0	0	0	4,625	2.62
December	1992	4,625	100	0	41	18	0	4,794	1.49
January	1993	4,794	100	0	0	0	0	5,006	2.33
February	1993	5,006	100	0	0	0	0	5,182	1.52

Hypothetical Portfolio Illustration

03-31-1991 to 03-31-2011

Investment Detail

Period		Beginning Balance	New Investment	Distribution/ Withdrawal	Total Reinvest	Charges & Fees	Taxes Due	Market Value	Total Return %
Totals		0	26,000	0	31,887	8,839	0	61,860	7.45
March	1993	5,182	100	0	72	20	0	5,393	2.14
April	1993	5,393	100	0	0	0	0	5,402	-1.67
May	1993	5,402	100	0	0	0	0	5,574	1.32
June	1993	5,574	100	0	44	22	0	5,738	1.15
July	1993	5,738	100	0	0	0	0	5,890	0.92
August	1993	5,890	100	0	0	0	0	6,184	3.29
September	1993	6,184	100	0	46	24	0	6,317	0.54
October	1993	6,317	100	0	0	0	0	6,495	1.23
November	1993	6,495	100	0	0	0	0	6,493	-1.57
December	1993	6,493	100	0	273	25	0	6,742	2.29
January	1994	6,742	100	0	0	0	0	7,048	3.06
February	1994	7,048	100	0	0	0	0	6,963	-2.62
March	1994	6,963	100	0	99	25	0	6,771	-4.20
April	1994	6,771	100	0	0	0	0	6,935	0.94
May	1994	6,935	100	0	0	0	0	7,074	0.56
June	1994	7,074	100	0	55	26	0	7,016	-2.22
July	1994	7,016	100	0	0	0	0	7,304	2.68
August	1994	7,304	100	0	0	0	0	7,622	2.98
September	1994	7,622	100	0	58	28	0	7,557	-2.16
October	1994	7,557	100	0	0	0	0	7,713	0.74
November	1994	7,713	100	0	0	0	0	7,557	-3.32
December	1994	7,557	100	0	172	29	0	7,686	0.38
January	1995	7,686	100	0	0	0	0	8,009	2.90
February	1995	8,009	100	0	0	0	0	8,366	3.20
March	1995	8,366	100	0	68	32	0	8,609	1.71
April	1995	8,609	100	0	0	0	0	8,910	2.34
May	1995	8,910	100	0	0	0	0	9,339	3.70
June	1995	9,339	100	0	67	36	0	9,582	1.52
July	1995	9,582	100	0	0	0	0	9,987	3.19
August	1995	9,987	100	0	0	0	0	10,135	0.49
September	1995	10,135	100	0	69	39	0	10,478	2.39
October	1995	10,478	100	0	0	0	0	10,495	-0.79
November	1995	10,495	100	0	0	0	0	11,064	4.47
December	1995	11,064	100	0	642	43	0	11,382	1.97
January	1996	11,382	100	0	0	0	0	11,699	1.90
February	1996	11,699	100	0	0	0	0	11,817	0.15
March	1996	11,817	100	0	182	45	0	12,037	1.02
April	1996	12,037	100	0	0	0	0	12,136	0.00
May	1996	12,136	100	0	0	0	0	12,444	1.71
June	1996	12,444	100	0	67	47	0	12,506	-0.30
July	1996	12,506	100	0	0	0	0	12,126	-3.84
August	1996	12,126	100	0	0	0	0	12,555	2.71
September	1996	12,555	100	0	68	49	0	13,045	3.11
October	1996	13,045	100	0	0	0	0	13,421	2.11

Hypothetical Portfolio Illustration

03-31-1991 to 03-31-2011

Investment Detail

Period		Beginning Balance	New Investment	Distribution/ Withdrawal	Total Reinvest	Charges & Fees	Taxes Due	Market Value	Total Return %
Totals		0	26,000	0	31,887	8,839	0	61,860	7.45
November	1996	13,421	100	0	0	0	0	14,355	6.21
December	1996	14,355	100	0	710	54	0	14,248	-1.44
January	1997	14,248	100	0	0	0	0	14,855	3.56
February	1997	14,855	100	0	0	0	0	15,104	1.00
March	1997	15,104	100	0	320	55	0	14,707	-3.29
April	1997	14,707	100	0	0	0	0	15,290	3.28
May	1997	15,290	100	0	0	0	0	16,273	5.78
June	1997	16,273	100	0	78	63	0	16,855	2.96
July	1997	16,855	100	0	0	0	0	18,163	7.17
August	1997	18,163	100	0	0	0	0	17,498	-4.21
September	1997	17,498	100	0	79	69	0	18,332	4.20
October	1997	18,332	100	0	0	0	0	17,977	-2.49
November	1997	17,977	100	0	0	0	0	18,420	1.91
December	1997	18,420	100	0	1,992	70	0	18,720	1.08
January	1998	18,720	100	0	0	0	0	18,923	0.55
February	1998	18,923	100	0	0	0	0	20,011	5.22
March	1998	20,011	100	0	292	79	0	20,979	4.34
April	1998	20,979	100	0	0	0	0	21,211	0.63
May	1998	21,211	100	0	0	0	0	21,258	-0.25
June	1998	21,258	100	0	94	81	0	21,613	1.20
July	1998	21,613	100	0	0	0	0	21,444	-1.25
August	1998	21,444	100	0	0	0	0	19,515	-9.46
September	1998	19,515	100	0	95	76	0	20,124	2.61
October	1998	20,124	100	0	0	0	0	21,208	4.89
November	1998	21,208	100	0	0	0	0	22,215	4.27
December	1998	22,215	100	0	1,226	86	0	22,849	2.41
January	1999	22,849	100	0	0	0	0	22,920	-0.13
February	1999	22,920	100	0	0	0	0	22,286	-3.20
March	1999	22,286	100	0	344	85	0	22,522	0.61
April	1999	22,522	100	0	0	0	0	23,820	5.32
May	1999	23,820	100	0	0	0	0	24,040	0.50
June	1999	24,040	100	0	106	93	0	24,787	2.69
July	1999	24,787	100	0	0	0	0	24,613	-1.10
August	1999	24,613	100	0	0	0	0	24,348	-1.49
September	1999	24,348	100	0	107	89	0	23,698	-3.08
October	1999	23,698	100	0	0	0	0	24,966	4.93
November	1999	24,966	100	0	0	0	0	25,467	1.61
December	1999	25,467	100	0	2,295	98	0	26,013	1.75
January	2000	26,013	100	0	0	0	0	25,772	-1.31
February	2000	25,772	100	0	0	0	0	26,043	0.66
March	2000	26,043	100	0	831	105	0	27,907	6.77
April	2000	27,907	100	0	0	0	0	28,007	0.00
May	2000	28,007	100	0	0	0	0	28,390	1.01
June	2000	28,390	100	0	160	107	0	28,399	-0.32

Hypothetical Portfolio Illustration

03-31-1991 to 03-31-2011

Investment Detail

Period		Beginning Balance	New Investment	Distribution/ Withdrawal	Total Reinvest	Charges & Fees	Taxes Due	Market Value	Total Return %
Totals		0	26,000	0	31,887	8,839	0	61,860	7.45
July	2000	28,399	100	0	0	0	0	28,821	1.13
August	2000	28,821	100	0	0	0	0	31,005	7.23
September	2000	31,005	100	0	162	117	0	31,189	0.27
October	2000	31,189	100	0	0	0	0	31,469	0.58
November	2000	31,469	100	0	0	0	0	30,879	-2.19
December	2000	30,879	100	0	2,642	121	0	32,098	3.63
January	2001	32,098	100	0	0	0	0	32,238	0.12
February	2001	32,238	100	0	0	0	0	31,300	-3.22
March	2001	31,300	100	0	842	115	0	30,427	-3.11
April	2001	30,427	100	0	0	0	0	31,639	3.65
May	2001	31,639	100	0	0	0	0	32,400	2.09
June	2001	32,400	100	0	228	121	0	32,110	-1.20
July	2001	32,110	100	0	0	0	0	32,752	1.69
August	2001	32,752	100	0	0	0	0	31,931	-2.81
September	2001	31,931	100	0	231	114	0	30,306	-5.40
October	2001	30,306	100	0	0	0	0	30,322	-0.28
November	2001	30,322	100	0	0	0	0	31,906	4.89
December	2001	31,906	100	0	392	121	0	32,111	0.33
January	2002	32,111	100	0	0	0	0	32,125	-0.27
February	2002	32,125	100	0	0	0	0	32,441	0.67
March	2002	32,441	100	0	470	125	0	33,318	2.40
April	2002	33,318	100	0	0	0	0	33,111	-0.92
May	2002	33,111	100	0	0	0	0	33,255	0.13
June	2002	33,255	100	0	185	119	0	31,650	-5.13
July	2002	31,650	100	0	0	0	0	29,886	-5.89
August	2002	29,886	100	0	0	0	0	30,297	1.04
September	2002	30,297	100	0	188	106	0	28,281	-6.98
October	2002	28,281	100	0	0	0	0	29,460	3.81
November	2002	29,460	100	0	0	0	0	30,508	3.22
December	2002	30,508	100	0	190	113	0	30,140	-1.53
January	2003	30,140	100	0	0	0	0	29,784	-1.51
February	2003	29,784	100	0	0	0	0	29,610	-0.92
March	2003	29,610	100	0	193	111	0	29,514	-0.66
April	2003	29,514	100	0	0	0	0	31,274	5.62
May	2003	31,274	100	0	0	0	0	33,224	5.92
June	2003	33,224	100	0	195	126	0	33,482	0.47
July	2003	33,482	100	0	0	0	0	33,768	0.56
August	2003	33,768	100	0	0	0	0	34,242	1.11
September	2003	34,242	100	0	197	129	0	34,221	-0.35
October	2003	34,221	100	0	0	0	0	35,405	3.17
November	2003	35,405	100	0	0	0	0	35,883	1.07
December	2003	35,883	100	0	199	142	0	37,612	4.54
January	2004	37,612	100	0	0	0	0	38,236	1.39
February	2004	38,236	100	0	0	0	0	39,099	2.00

Hypothetical Portfolio Illustration

03-31-1991 to 03-31-2011

Investment Detail

Period		Beginning Balance	New Investment	Distribution/ Withdrawal	Total Reinvest	Charges & Fees	Taxes Due	Market Value	Total Return %
Totals		0	26,000	0	31,887	8,839	0	61,860	7.45
March	2004	39,099	100	0	201	145	0	38,537	-1.69
April	2004	38,537	100	0	0	0	0	38,060	-1.50
May	2004	38,060	100	0	0	0	0	38,305	0.38
June	2004	38,305	100	0	203	147	0	38,946	1.41
July	2004	38,946	100	0	0	0	0	38,561	-1.25
August	2004	38,561	100	0	0	0	0	38,806	0.38
September	2004	38,806	100	0	205	148	0	39,402	1.28
October	2004	39,402	100	0	0	0	0	39,943	1.12
November	2004	39,943	100	0	0	0	0	41,318	3.19
December	2004	41,318	100	0	346	161	0	42,696	3.09
January	2005	42,696	100	0	0	0	0	42,102	-1.62
February	2005	42,102	100	0	0	0	0	43,244	2.48
March	2005	43,244	100	0	779	161	0	42,666	-1.57
April	2005	42,666	100	0	0	0	0	42,563	-0.48
May	2005	42,563	100	0	0	0	0	43,424	1.79
June	2005	43,424	100	0	231	165	0	43,794	0.62
July	2005	43,794	100	0	0	0	0	45,068	2.68
August	2005	45,068	100	0	0	0	0	45,475	0.68
September	2005	45,475	100	0	233	173	0	45,943	0.81
October	2005	45,943	100	0	0	0	0	45,580	-1.01
November	2005	45,580	100	0	0	0	0	46,350	1.47
December	2005	46,350	100	0	1,826	175	0	46,587	0.30
January	2006	46,587	100	0	0	0	0	47,546	1.84
February	2006	47,546	100	0	0	0	0	47,646	0.00
March	2006	47,646	100	0	1,005	180	0	47,759	0.03
April	2006	47,759	100	0	0	0	0	48,629	1.61
May	2006	48,629	100	0	0	0	0	48,012	-1.47
June	2006	48,012	100	0	273	180	0	47,881	-0.48
July	2006	47,881	100	0	0	0	0	48,812	1.74
August	2006	48,812	100	0	0	0	0	49,467	1.14
September	2006	49,467	100	0	284	190	0	50,388	1.66
October	2006	50,388	100	0	0	0	0	51,438	1.88
November	2006	51,438	100	0	0	0	0	52,153	1.20
December	2006	52,153	100	0	1,856	199	0	52,913	1.27
January	2007	52,913	100	0	0	0	0	53,419	0.77
February	2007	53,419	100	0	0	0	0	52,938	-1.09
March	2007	52,938	100	0	528	201	0	53,318	0.53
April	2007	53,318	100	0	0	0	0	55,473	3.85
May	2007	55,473	100	0	0	0	0	56,984	2.54
June	2007	56,984	100	0	324	211	0	56,136	-1.66
July	2007	56,136	100	0	0	0	0	54,580	-2.95
August	2007	54,580	100	0	0	0	0	55,153	0.87
September	2007	55,153	100	0	327	211	0	56,087	1.51
October	2007	56,087	100	0	0	0	0	56,843	1.17

Hypothetical Portfolio Illustration

03-31-1991 to 03-31-2011

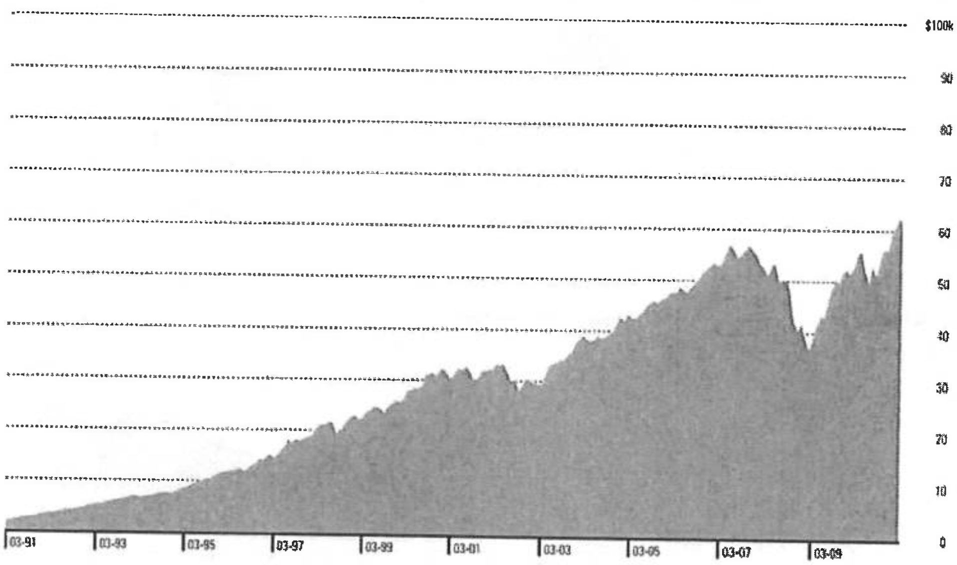
Investment Detail

Period		Beginning Balance	New Investment	Distribution/ Withdrawal	Total Reinvest	Charges & Fees	Taxes Due	Market Value	Total Return %
Totals		0	26,000	0	31,887	8,839	0	61,860	7.45
November	2007	56,843	100	0	0	0	0	55,987	-1.68
December	2007	55,987	100	0	2,260	207	0	55,007	-1.93
January	2008	55,007	100	0	0	0	0	53,302	-3.28
February	2008	53,302	100	0	0	0	0	52,404	-1.87
March	2008	52,404	100	0	408	193	0	51,221	-2.45
April	2008	51,221	100	0	0	0	0	53,081	3.44
May	2008	53,081	100	0	0	0	0	53,307	0.24
June	2008	53,307	100	0	344	188	0	50,022	-6.35
July	2008	50,022	100	0	0	0	0	49,931	-0.38
August	2008	49,931	100	0	0	0	0	50,476	0.89
September	2008	50,476	100	0	348	180	0	47,802	-5.49
October	2008	47,802	100	0	0	0	0	42,335	-11.65
November	2008	42,335	100	0	0	0	0	40,510	-4.55
December	2008	40,510	100	0	351	157	0	41,778	2.88
January	2009	41,778	100	0	0	0	0	39,158	-6.51
February	2009	39,158	100	0	0	0	0	36,790	-6.30
March	2009	36,790	100	0	274	146	0	38,860	5.36
April	2009	38,860	100	0	0	0	0	41,319	6.07
May	2009	41,319	100	0	0	0	0	43,324	4.61
June	2009	43,324	100	0	277	163	0	43,205	-0.51
July	2009	43,205	100	0	0	0	0	46,415	7.20
August	2009	46,415	100	0	0	0	0	48,437	4.14
September	2009	48,437	100	0	213	191	0	50,626	4.31
October	2009	50,626	100	0	0	0	0	49,726	-1.97
November	2009	49,726	100	0	0	0	0	51,962	4.30
December	2009	51,962	100	0	215	197	0	52,216	0.30
January	2010	52,216	100	0	0	0	0	51,712	-1.16
February	2010	51,712	100	0	0	0	0	52,887	2.08
March	2010	52,887	100	0	269	208	0	55,340	4.45
April	2010	55,340	100	0	0	0	0	55,912	0.85
May	2010	55,912	100	0	0	0	0	52,226	-6.77
June	2010	52,226	100	0	237	188	0	50,001	-4.45
July	2010	50,001	100	0	0	0	0	53,022	5.84
August	2010	53,022	100	0	0	0	0	51,352	-3.34
September	2010	51,352	100	0	240	205	0	54,355	5.65
October	2010	54,355	100	0	0	0	0	56,232	3.27
November	2010	56,232	100	0	0	0	0	56,058	-0.49
December	2010	56,058	100	0	287	223	0	59,111	5.27
January	2011	59,111	100	0	0	0	0	60,587	2.33
February	2011	60,587	100	0	0	0	0	62,479	2.96
March	2011	62,479	100	0	266	233	0	61,860	-1.15

Hypothetical Portfolio Illustration Continued

03-31-1991 to 03-31-2011

Security Summary



% of Total Market Value

Start 03-31-1991

End 03-31-2011

Investment Assumptions

Investment Name	Holding Period Start	Holding Period End	Initial Investment Amount	Subsequent Invest/Withdwl Amount	Subsequent Invest/Withdwl Freq	Reinvest Distributions Income	Reinvest Distributions Cap Gains	Liquidate	Rebalance %	Charges and Fees Front Load	Charges and Fees Annual Fee%	Charges and Fees Deferred Load Amount%	Charges and Fees Period Years	Market Value End $
Invesco Van Kampen Equity and Income A (USD)	03-91	03-11	2,000	100	Mon	Y	Y	N	—	$0.00	1.50	0.00-0.00	—	61,860

Release date 03-31-2011

Portfolio Snapshot

Portfolio Value
$61,860

Benchmark
S&P 500 TR (USD)

Analysis 03-31-2011

Asset Allocation	Portfolio Net %	Bmark Net %
Cash	2.94	0.00
US Stocks	57.78	99.90
Non-US Stocks	4.79	0.10
Bonds	19.24	0.00
Other/Not Clsfd	**15.25**	**0.00**

Morningstar Equity Style Box %

Total Stock Holdings 78

Value	Core	Growth	
41	36	17	Large
3	2	1	Mid
0	0	0	Small

0-10 10-25 25-50 >50

Morningstar Fixed Income Style Box %

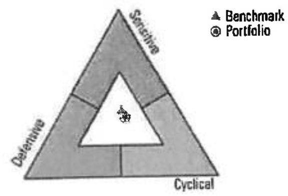
Total Bond Holdings 199

Ltd	Mod	Ext	
0	0	0	High
0	100	0	Med
0	0	0	Low

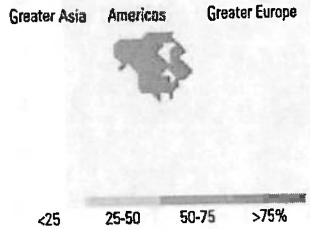
0-10 10-25 25-50 >50

Stock Analysis 03-31-2011

Stock Sectors

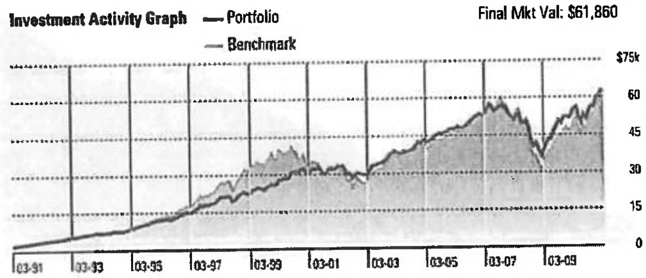


World Regions



Performance 03-31-2011

Investment Activity Graph



% of Stocks	Portfolio %	Bmark %
Cyclical	**36.43**	**28.50**
Basic Matls	3.28	2.96
Consumer Cycl	11.21	9.25
Financial Svs	21.94	14.67
Real Estate	0.00	1.62
Sensitive	**39.10**	**46.85**
Commun Svs	6.10	4.25
Energy	14.79	13.01
Industrials	9.87	12.93
Technology	8.34	16.66
Defensive	**24.47**	**24.65**
Consumer Def	10.71	10.73
Healthcare	9.71	10.79
Utilities	4.05	3.13
Not Classified	**0.00**	**0.00**

% of Stocks	Portfolio %	Bmark %
Greater Europe	**6.44**	**0.10**
United Kingdom	3.79	0.00
Europe-Developed	2.65	0.10
Europe-Emerging	0.00	0.00
Africa/Middle East	0.00	0.00
Americas	**92.35**	**99.91**
North America	92.35	99.91
Latin America	0.00	0.00
Greater Asia	**1.21**	**0.00**
Japan	1.21	0.00
Australasia	0.00	0.00
Asia-Developed	0.00	0.00
Asia-Emerging	0.00	0.00
Not Classified	**0.00**	**0.00**

Trailing Returns	3 Mo	1 Yr	3 Yr	5 Yr	10 Yr
Portfolio Return	4.14	9.52	4.25	3.01	4.41
Benchmark Return	5.92	15.81	2.88	2.87	3.58
+/- Benchmark Return	-1.78	-6.29	1.37	0.14	0.83

Best/Worst Time Periods	Best %	Worst %
3 Months	16.90 (Mar 09-May 09)	-20.30 (Sep 08-Nov 08)
1 Year	39.93 (Mar 09-Feb 10)	-31.73 (Mar 08-Feb 09)
3 Years	23.01 (Apr 95-Mar 98)	-10.87 (Mar 06-Feb 09)

Portfolio Yield (03-31-2011)	Yield %
12-Month Yield	1.68

Performance Disclosure

The performance data quoted represents past performance and does not guarantee future results. The investment return and principal value of an investment will fluctuate thus an investor's shares, when redeemed, may be worth more or less than their original cost. Current performance may be lower or higher than return data quoted herein. For performance data current to the most recent month-end, please visit http://advisor.morningstar.com/familyinfo.asp.

See Disclosure Page for Standardized Returns.

Holdings 03-31-2011

Top 1 holding out of 1

	Symbol	Type	Holding Value $	% Assets
Invesco Van Kampen Equity and Income A (USD)	ACEIX	MF	61,860	100.00

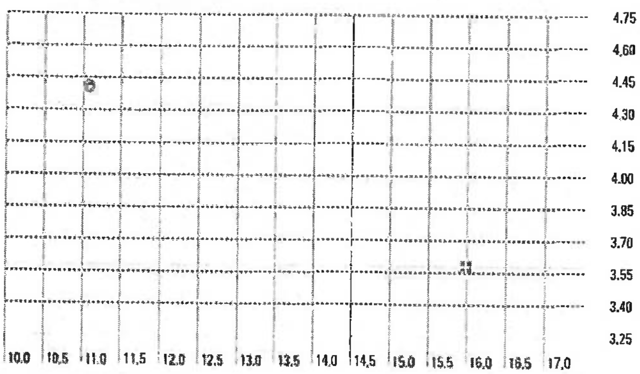
Portfolio Snapshot

Portfolio Value
$61,860

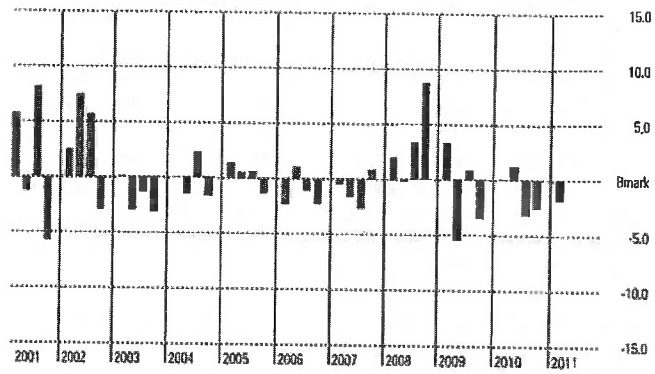
Benchmark
S&P 500 TR (USD)

Risk Analysis 03-31-2011

Risk/Reward Scatterplot ◉ Portfolio ● Holdings ▪ Bmark 10 Year Mean



Performance History Graph ■ Portfolio Quarterly returns +/- Benchmark in %



Risk and Return Statistics	3 Yr Portfolio	3 Yr Bmark	5 Yr Portfolio	5 Yr Bmark	10 Yr Portfolio	10 Yr Bmark
Standard Deviation	15.97	21.89	12.99	17.87	11.07	15.98
Mean	4.25	2.88	3.01	2.87	4.41	3.58
Sharpe Ratio	0.30	0.20	0.12	0.12	0.25	0.15

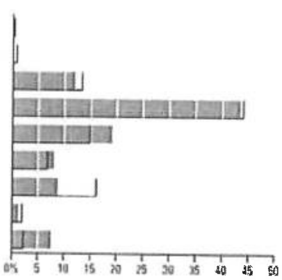
MPT Statistics	3 Yr Portfolio	5 Yr Portfolio	10 Yr Portfolio
Alpha	1.61	0.06	1.13
Beta	0.72	0.71	0.66
R-Squared	96	95	91

Fundamental Analysis 03-31-2011

Market Maturity % of Stocks	Portfolio	Bmark
Developed Markets	100.00	100.00
Emerging Markets	0.00	0.00
Not Available	**0.00**	**0.00**

Geometric Avg Capitalization ($Mil)	
Portfolio	44,879.69
Benchmark	50,179.19

Type Weightings % of Stocks	■ Portfolio	┘ Bmark
High Yield	0.00	0.23
Distressed	0.00	0.67
Hard Asset	11.94	13.30
Cyclical	43.35	43.95
Slow Growth	19.06	14.82
Classic Growth	8.02	6.74
Aggressive Growth	8.80	16.09
Speculative Growth	1.23	1.98
Not Available	**7.62**	**2.22**



Valuation Multiples	Portfolio	Bmark
Price/Earnings	11.06	16.13
Price/Book	1.73	2.26
Price/Sales	1.12	1.39
Price/Cash Flow	7.68	9.50

Profitability % of Stocks	Portfolio 2011-03	Bmark 2011-03
Net Margin	11.19	12.90
ROE	15.50	20.92
ROA	5.86	8.47
Debt/Capital	38.03	35.67

Fund Statistics	
Potential Cap Gains Exposure	7.85
Avg Net Expense Ratio	0.78
Avg Gross Expense Ratio	0.78

Credit Quality Breakdown	% of Bonds
AAA	55.11
AA	10.96
A	13.73
BBB	20.20
BB	0.00
B	0.00
Below B	0.00
NR/NA	0.00

Interest Rate Risk	Portfolio
Avg Eff Maturity	7.40
Avg Eff Duration (total portfolio)	4.85
Avg Credit Quality	—
Avg Wtd Coupon	3.55

Portfolio Snapshot

Portfolio Value
$61,860

Benchmark
S&P 500 TR (USD)

Standardized and Tax Adjusted Returns

The performance data quoted represents past performance and does not guarantee future results. The investment return and principal value of an investment will fluctuate thus an investor's shares, when redeemed, may be worth more or less than their original cost. Current performance may be lower or higher than return data quoted herein. For performance data current to the most recent month-end please visit http://advisor.morningstar.com/familyinfo.asp

An investment in a money-market vehicle is not insured or guaranteed by the FDIC or any other government agency. The current yield quotation reflects the current earnings of the money market more closely than the total return quotation. Although money markets seek to preserve the value of your investment at $1.00 per share, it is possible to lose money by investing in them.

Standardized Returns assume reinvestment of dividends and capital gains. It depicts performance without adjusting for the effects of taxation, but are adjusted to reflect sales charges and ongoing fund expenses.

If adjusted for taxation, the performance quoted would be significantly reduced.

For variable annuities, additional expenses will be taken in account, including M&E risk charges, fund-level expenses such as management fees and operating fees, and contract-level administration fees, charges such as surrender, contract and sales charges.

After-tax returns are calculated using the highest individual federal marginal income tax rates, and do not reflect the impact of state and local taxes. Actual after tax returns depend on the investor's tax situation and may differ from those shown. The after tax returns shown are not relevant to investors who hold their fund shares through tax-deferred arrangements such as 401(k) plans or an IRA. After-tax returns exclude the effects of either the alternative minimum tax or phase-out of certain tax credits. Any taxes due are as of the time the distributions are made, and the taxable amount and tax character of each distribution is as specified by the fund on the dividend declaration date. Due to foreign tax credits or realized capital losses, after-tax returns may be greater than before tax returns. After-tax returns for exchange-traded funds are based on net asset value.

Annualized returns 03-31-2011

Standardized Returns (%)	7-day Yield	1Yr	5Yr	10Yr	Since Inception	Inception Date	Max Front Load %	Max Back Load %	Net Exp Ratio %	Gross Exp Ratio %
Invesco Van Kampen Equity and Income A (USD)	—	4.96	3.23	5.31	10.30	08-03-1960	**5.50**	**NA**	**0.78**	**0.78**
BarCap US Agg Bond TR USD	—	**5.12**	**6.03**	**5.56**	—	—				
MSCI EAFE NR USD	—	**10.42**	**1.30**	**5.39**	—	—				
S&P 500 TR	—	**15.65**	**2.62**	**3.29**	—	—				
USTREAS T-Bill Auction Ave 3 Mon	—	**0.15**	**2.08**	**2.14**	—	—				

Return after Tax (%)	On Distribution					On Distribution and Sales of Shares			
	1Yr	5Yr	10Yr	Since Inception	Inception Date	1Yr	5Yr	10Yr	Since Inception
Invesco Van Kampen Equity and Income A (USD)	4.28	2.14	4.18	6.37	08-03-1960	3.18	2.20	3.99	6.27

Portfolio Snapshot

Portfolio Value
$61,860

Benchmark
S&P 500 TR (USD)

Illustration Returns

Total 1 holding as of 03-31-2011	Symbol	Type	Holdings Date	% of Assets	Holding Value $	7-day Yield	1 Yr Ret %	3 Yr Ret %	5 Yr Ret %	10 Yr Ret %
Invesco Van Kampen Equity and Income A (USD)	ACEIX	MF	12-2010	100.00	61,860	—	9.52	4.25	3.01	4.41

Performance Disclosure

The performance data quoted represents past performance and does not guarantee future results. The investment return and principal value of an investment will fluctuate thus an investor's shares, when redeemed, may be worth more or less than their original cost. Current performance may be lower or higher than return data quoted herein. For performance data current to the most recent month-end, please visit http://advisor.morningstar.com/familyinfo.asp.

See Disclosure Page for Standardized Returns.

Hypothetical Report

Disclosure Statement

General

This is an illustration of a simulated investment and assumes the portfolio holding(s) were purchased on the first day of the period indicated. Sales and tax charges, including those required in the event of transfers between assets, are taken into account at the rates shown and may be higher or lower than what an investor would have actually paid had the investments been purchased then or now. The performance data represents past performance and is not indicative of future results. Principal value and investment returns will fluctuate, and an investor's shares/units when redeemed may be worth more or less than the original investment.

The underlying holdings of the portfolio are not federally or FDIC-insured and are not deposits or obligations of, or guaranteed by, any financial institution. Investment in securities involve investment risks including possible loss of principal and fluctuation in value.

The investment returns do not reflect active trading and do not necessarily reflect the results that might have been achieved by active management of the account. The investment returns of other clients of the adviser may differ materially from the investment portrayed.

The information contained in this report is from the most recent information available to Morningstar as of the release date, and may or may not be an accurate reflection of the current composition of the securities included in the portfolio. There is no assurance that the weightings, composition and ratios will remain the same.

Pre-inception Returns

The analysis in this report may be based, in part, on adjusted historical returns for periods prior to the fund's actual inception. These calculated returns reflect the historical performance of the oldest share class of the fund, adjusted to reflect the fees and expenses of this share class. These fees and expenses are referenced in the report's Charges and Fees section.

When pre-inception data are presented in the report, the header at the top of the report will indicate this.

While the inclusion of pre-inception data provides valuable insight into the probable long-term behavior of newer share classes of a fund, investors should be aware that an adjusted historical return can only provide an approximation of that behavior. For example, the fee structures between a retail share class will vary from that of an institutional share class, as retail shares tend to have higher operating expenses and sales charges. These adjusted historical returns are not actual returns. Calculation methodologies utilized by Morningstar may differ from those applied by other entities, including the fund itself.

The investment returns do not necessarily reflect the deduction of all investment advisory fees. Client investment returns may be reduced if additional fees are incurred.

Performance for closed-end and exchange-traded funds is calculated based on the fund's end of the day market prices as reported by the New York Stock Exchange. Separate account performance is based on the mean experience of an investor in the account.

This illustration may reflect the results of systematic investments and/or withdrawals. Systematic investment does not ensure a profit, nor does it protect the investor against a loss in a declining market. Also, systematic investing will not keep an investor from losing money if shares are sold when the market is down.

Investment Summary Graph

The investment summary graph plots the approximate market value of the security or portfolio over the investing horizon. It may also include the total investment assumed in the illustration and/or a benchmark. Total investment includes dollar inflows and outflows, including inflows representing noted taxes and annual fees paid out of pocket. If a benchmark index is included on a graph, it assumes a similar pattern of investment/withdrawal as for the security or portfolio. Taxes and transaction costs are aslo applied to the benchmark index. Note that direct investment in an index is not possible. Indexes are unmanaged portfolios representing different asset classes, with varying levels of associated risk. The benchmark index included in the graph may or may not represent an appropriate or accurate comparison with the security or portfolio illustrated.

Standardized Returns

For ETFs, the standardized returns reflect performance, both at market price and NAV price, without adjusting for the effects of taxation or brokers commissions. These returns are adjusted to reflect all ongoing ETF expenses and assume reinvestment of dividends and capital gains. If adjusted, the effects of taxation would reduce the performance quoted.

For HOLDs, the standardized returns reflect performance at market price, without adjusting for the effects of taxation or brokers commissions. These returns are adjusted to reflect all ongoing expenses and assume reinvestment of dividends and capital gains. If adjusted, the effects of taxation would reduce the performaance quoted.

For money market mutual funds, standardized return is total return adjusted for sales charges and reflects all ongoing fund expenses. Current 7-day yield more closely reflects the current earnings of the money market fund than the total return quotation.

For mutual funds, standardized return is total return adjusted for sales charges and reflects all ongoing fund expenses. Following this disclosure statement, standardized returns for each portfolio holding are shown.

For VA subaccounts, standardized return is total return based on its inception date within the separate account and is adjusted to reflect recurring and non-recurring charges such as surrender fees, contract charges, maximum front-end load, maximum deferred load, maximum M&E risk charge, administration fees, and actual ongoing fund-level expenses.

For VL subaccounts, standardized return is total return based on its inception date within the separate account and is adjusted to reflect recurring and non-recurring charges such as surrender fees, contract charges, maximum front-end load, maximum deferred load, maximum M&E risk charge, administration fees, and actual ongoing fund-level expenses. For VLs, additional fees specific to a VL policy such as transfer fees and cost of insurance fees, which are based on specific characteristics on an individual, are not included. If VL fees were included in the return calculations, the performance would have been significantly lower. An investor should contact their financial advisor and ask for a personalized performance illustration, either hypothetical or historical, which reflects all applicable fees and charges including the cost of insurance. Please review the prospectus and SAI for more detailed information.

Definitions of Report Terms

*Annual Fee Paid:*Your advisor was able to specify whether annual fees, if any, should be assumed paid out of pocket or from selling shares of securities held in the illustration.

*Average Annualized Return:*Average annualized money-weighted return (internal rate of return). In illustrations with time periods less than one year, this figure is not annualized.

*Capital Gains (Individual Report):*Percentage of the total market value of the holding that is attributable to the reinvestment of capital gains distributions.

*Charges & Fees (Investment Detail):*The sum of fees charged to the investor during the period, including front or deferred loads, VA charges, and annual fees.

*Cumulative Return:*The total money-weighted return of the investment over the entire time period of the illustration.

*Distribution/Withdrl:*The sum of distributions not reinvested, plus any cash withdrawals during the period.

*Income (Individual Report):*The percentage of the total market value of the holding that is attributable to the reinvestment of income or dividend distributions.

*Liquidate:*Indicates whether the advisor chose that the holding be liquidated on the end date.

*Median (Comparison Report):*The total money-weighted return (internal rate of return) of the median security in the illustration for the calendar year indicated.

*New investment:*Any new cash invested during the period.

*Principal (Individual Reports):*The percentage of the total market value of the holding that is attributable to new investment.

*Rebalance(Planning Assumptions):*Indicates whether rebalancing is used, and its frequency. "No" indicates no rebalancing. Options for rebalancing frequency are monthly, quarterly, semi-annually, and annually.

*Rebalance(Investment Assumptions):*Percentage of total asset allocation to be maintained in this holding through rebalancing.

*Securities Returns(Comparison Report):*The total money-weighted return (internal rate of return) for the holding in the calendar year indicated, taking into account cash flows, charges, and fees.

*Subsequent Invest/Withdwl:*The amount, type, and frequency of subsequent investments or withdrawals from the holding. Withdrawals are represented by a negative number. Systematic investments and withdrawals may be made monthly, quarterly, semi-annually, or annually. If "Custom, " a custom schedule of investments or withdrawals was used.

*Taxes Due:*The total amount of taxes due from the investor, determined by applying specified tax rates to distributions and sale of shares during each calendar year.

*Taxes Paid:*Your advisor was able to specify whether taxes, if any, should be assumed paid out of pocket or from selling shares of securities held in the illustration.

*Net Dollars Invested:*The total out-of-pocket expense for the investor. Includes new investment, annual fees paid to advisor, and taxes due. This figure is net of withdrawals, including liquidation.

*Total Reinvest:*The sum of distributions reinvested during the period.

*Total Return %:*The total money-weighted return (internal rate of return) on investments for the period.

Portfolio Snapshot Report Disclosure Statement

General

Investment portfolios illustrated in this report can be scheduled or unscheduled. With an unscheduled portfolio, the user inputs only the portfolio holdings and their current allocations. Morningstar calculates returns using the given allocations assuming monthly rebalancing. Taxes, loads, and sales charges are not taken into account.

With "scheduled" portfolios, users input the date and amount for all investments into and withdrawals from each holding, as well as tax rates, loads, and other factors that would have affected portfolio performance. A hypothetical illustration is one type of scheduled portfolio.

Both scheduled and unscheduled portfolios are theoretical, for illustrative purposes only, and are not reflective of an investors actual experience. For both scheduled and unscheduled portfolios, the performance data given represents past performance and should not be considered indicative of future results. Principal value and investment return of stocks, mutual funds, and variable annuity/life products will fluctuate, and an investor's shares/units when redeemed will be worth more or less than the original investment. Stocks, mutual funds, and variable annuity/life products are not FDIC-insured, may lose value, and are not guaranteed by a bank or other financial institution. Portfolio statistics change over time.

Used as supplemental sales literature, the Portfolio Snapshot report must be preceded or accompanied by the fund/policy's current prospectus or equivalent. In all cases, this disclosure statement should accompany the Portfolio Snapshot report. Morningstar is not itself a FINRA-member firm.

The underlying holdings of the portfolio are not federally or FDIC-insured and are not deposits or obligations of, or guaranteed by any financial institution. Investment in securities involve investment risks including possible loss of principal and fluctuation in value.

The information contained in this report is from the most recent information available to Morningstar as of the release date, and may or may not be an accurate reflection of the current composition of the securities included in the portfolio. There is no assurance that the weightings, composition and ratios will remain the same.

Items to Note Regarding Certain Underlying Securities

A closed-end fund is an investment company, which typically makes one public offering of a fixed number of shares. Thereafter, shares are traded on a secondary market such as the New York Stock Exchange. As a result, the secondary market price may be higher or lower than the closed-end fund's net asset value (NAV). If these shares trade at a price above their NAV, they are said to be trading at a premium. Conversely, if they are trading at a price below their NAV, they are said to be trading at a discount.

An exchange-traded fund (ETF) is an investment company that typically has an investment objective of striving to achieve a similar return as a particular market index. The ETF will invest in either all or a representative sample of the securities included in the index it is seeking to imitate. Like closed-end funds, ETFs can be traded on a secondary market and thus have a market price that may be higher or lower than its net asset value. If these shares trade at a price above their NAV, they are said to be trading at a premium. Conversely, if they are trading at a price below their NAV, they are said to be trading at a discount.

A money market fund is an investment company that invests in commercial paper, banker's acceptances, repurchase agreements, government securities, certificates of deposit and other highly liquid securities, and pays money market rates of interest. Money markets are not FDIC-insured, may lose money, and are not guaranteed by a bank or other financial institution. Although the money market seeks to preserve a stable per share value (i.e. $1.00 per share), it is possible to lose money by investment in the fund.

Unit investment trust (UIT) is an investment company organized under a trust agreement between a sponsor and trustee. UITs typically purchase a fixed portfolio of securities and then sell units in the trust to investors. The major difference between a UIT and a mutual fund is that a mutual fund is actively managed, while a UIT is not. On a periodic basis, UITs usually distribute to the unit holder their pro rata share of the trust's net investment income and net realized capital gains, if any. If the trust is one that invests only in tax-free securities, then the income from the trust is also tax-free. UITs generally make one public offering of a fixed number of units. However, in some cases, the sponsor will maintain a secondary market that allows existing unit holders to sell their units and for new investors to buy units.

Variable annuities are tax-deferred investments structured to convert a sum of money into a series of payments over time. Variable annuity policies have limitations and are not viewed as short-term liquid investments. An insurance company's fulfillment of a commitment to pay a minimum death benefit, a schedule of payments, a fixed investment account guaranteed by the insurance company, or another form of guarantee depends on the claims-paying ability of the issuing insurance company. Any such guarantee does not affect or apply to the investment return or principal value of the separate account and its subaccount. The financial ratings quoted for an insurance company do not apply to the separate account and its subaccount. If the variable annuity subaccount is invested in a money-market fund, although it seeks to preserve a stable per share value (i.e. $1.00 per share), it is possible to lose money by investment in the fund.

Variable life insurance is a cash-value life insurance that has a variable cashvalue and/or death benefit depending on the investment performance of the subaccount into which premium payments are invested. Unlike traditional life insurance, variable life insurance has inherent risks associated with it, including market volatility, and is not viewed as a short-term liquid investment. For more information on a variable life product, including each subaccount, please read the current prospectus. Please note, the financial ratings noted on the report are quoted for an insurance company and do not apply to the separate account and its subaccount. If the variable life subaccount is invested in a money-market fund, although it seeks to preserve a stable per share value (i.e. $1.00 per share), it is possible to lose money by investment in the fund.

Pre-inception Returns

The analysis in this report may be based, in part, on adjusted historical returns for periods prior to the fund's actual inception. These calculated returns reflect the historical performance of the oldest share class of the fund, adjusted to reflect the fees and expenses of this share class. These fees and expenses are referenced in the report's list of holdings and again on the standardized returns page. When pre-inception data is presented in the report, the header at the top of the report will indicate this and the affected data elements will be displayed in italics.

While the inclusion of pre-inception data provides valuable insight into the probable long-term behavior of newer share classes of a fund, investors should be aware that an adjusted historical return can only provide an approximation of that behavior. For example, the fee structures between a retail share class will vary from that of an institutional share class, as retail shares tend to have higher operating expenses and sales charges. These adjusted historical returns are not actual returns. Calculation methodologies utilized by Morningstar may differ from those applied by other entities, including the fund itself.

Scheduled Portfolio Trailing Returns

Scheduled Portfolios are customized by the user to account for loads, taxes, cash flows and specific investment dates. Scheduled portfolios use the portfolio's investment history to calculate final market values and returns. For scheduled portfolios, both individual holdings and portfolio returns are internal-rate-of-return calculations that reflect the timing and dollar size of all purchases and sales. For stocks and mutual funds, sales charges and tax rates are taken into account as specified by the user (except in the pre-tax returns, which reflect the impact of sales charges but not taxes). Note that in some scheduled portfolio illustrations, dividends and capital gains distributions, if applicable, are reinvested at the end of the month in which they are made at the month-end closing price. This can cause discrepancies between calculated returns and actual investor experience.

Scheduled Portfolio Returns-Based Performance Data

For scheduled portfolios, the monthly returns used to calculate alphas, betas, R-squareds, standard deviations, Sharpe ratios and best/worst time-period data are internal rates of return.

Important VA Disclosure for Scheduled Portfolios

For variable annuity products, policy level charges (other than front-end loads, if input by the advisor) are not factored into returns. When withdrawals and liquidations are made, increases in value over the purchase price are taxed at the capital gains rate that currently is in effect. This is not reflective of the actual tax treatment for these products, which requires the entire withdrawal to be taxed at the income tax rate. If adjusted for sales charges and the effects of taxation, the subaccount returns would be reduced.

Scheduled Portfolio Investment Activity Graph

The historic portfolio values that are graphed are those used to track theportfolio when calculating returns.

Unscheduled Portfolio Returns

Monthly total returns for unscheduled portfolios are calculated by applying theending period holding weightings supplied by the user to an individual holding's monthly returns. When monthly returns are unavailable for a holding (ie. Due to it not being in existence during the historical period being reported), the remaining portfolio holdings are re-weighted to maintain consistent proportions. Inception dates are listed in the Disclosure for Standardized and Tax Adjusted Returns. Trailing returns are calculated by geometrically linking these weighted-average monthly returns. Unscheduled portfolio returns thus assume monthly rebalancing. Returns for individual holdings are simple time-weighted trailing returns. Neither portfolio returns nor holding returns are adjusted for loads or taxes, and if adjusted for, would reduce the returns stated. The returns stated assume the reinvestment of dividends and capital gains. Mutual fund returns include all ongoing fund expenses. VA/VL returns reflect subaccount level fund expenses, including M&E expenses, administration fees, and actual ongoing fund level expenses.

Unscheduled Portfolio Investment Activity Graph

The historic performance data graphed is extrapolated from the ending portfolio value based on monthly returns.

Benchmark Returns

Benchmark returns may or may not be adjusted to reflect ongoing expenses such as sales charges. An investment's portfolio may differ significantly from the securities in the benchmark.

Returns for custom benchmarks are calculated by applying user-supplied weightings to each benchmark's returns every month. Trailing returns are calculated by geometrically linking these weighted-average monthly returns. Custom benchmark returns thus assume monthly rebalancing.

Standardized Returns

For mutual funds, standardized return is total return adjusted for sales charges, and reflects all ongoing fund expenses. Following this disclosure statement, standardized returns for each portfolio holding are shown.

For money market mutual funds, standardized return is total return adjusted for sales charges and reflects all ongoing fund expenses. Current 7-day yield more closely reflects the current earnings of the money market fund than the total return quotation.

For VA subaccounts, standardized return is total return based on its inception date within the separate account and is adjusted to reflect recurring and non-recurring charges such as surrender fees, contract charges, maximum front-end load, maximum deferred load, maximum M&E risk charge, administration fees and actual ongoing fund-level expenses.

For ETFs, the standardized returns reflect performance, both at market priceand NAV price, without adjusting for the effects of taxation or brokers commissions. These returns are adjusted to reflect all ongoing ETF expenses and assume reinvestment of dividends and capital gains. If adjusted, the effects of taxation would reduce the performance quoted.

The charges and expenses used in the standardized returns are obtained from the most recent prospectus and/or shareholder report available to Morningstar. For mutual funds and VAs, all dividends and capital gains are assumed to be reinvested. For stocks, stock acquired via divestitures is assumed to be liquidated and reinvested in the original holding.

Non-Standardized Returns

For mutual funds, total return is not adjusted for sales charges and reflects all ongoing fund expenses for various time periods. These returns assume reinvestment of dividends and capital gains. If adjusted for sales charges and the effects of taxation, the mutual fund returns would be reduced. Please note these returns can include pre-inception data and if included, this data will be represented in italics.

For money market funds, total return is not adjusted for sales charges and reflects all ongoing fund expenses for various time periods. These returns assume reinvestment of dividends and capital gains. If adjusted for sales charges and the effects of taxation, the money market returns would be reduced.

For VA and VL subaccounts, non-standardized returns illustrate performance that is adjusted to reflect recurring and non-recurring charges such as surrender fees, contract charges, maximum front-end load, maximum deferred load, maximum M&E risk charge, administrative fees and underlying fund-level expenses for various time periods. Non-Standardized performance returns assume reinvestment of dividends and capital gains. If adjusted for the effects of taxation, the subaccount returns would be significantly reduced. Please note these returns can include pre-inception data and if included, this data will be represented in italics.

Investment Advisory Fees

The investment(s) returns do not necessarily reflect the deduction of all investment advisory fees. Client investment returns will be reduced if additional advisory fees are incurred such as deferred loads, redemption fees, wrap fees, or other account charges.

Investment Style

The Morningstar Style Box reveals a fund's investment style as of the date noted on this report.

For equity funds the vertical axis shows the market capitalization of the long stocks owned and the horizontal axis shows investment style (value, blend, or growth).

For fixed-income funds, the vertical axis shows the credit quality of the long bonds owned and the horizontal axis shows interest rate sensitivity as measured by a bond's effective duration.

Morningstar seeks credit rating information from fund companies on a periodic basis (e.g., quarterly). In compiling credit rating information, Morningstar instructs fund companies to only use ratings that have been assigned by a Nationally Recognized Statistical Rating Organization (NRSRO). If two NRSROs have rated a security, fund companies are to report the lowest rating; if three or more NRSROs have rated the same security differently, fund companies are to report the rating that is in the middle. For example, if NRSRO X rates a security AA-, NRSRO Y rates the same security an A and NRSRO Z rates it a BBB+, the fund company should use the credit rating of 'A' in its reporting to Morningstar. PLEASE NOTE: Morningstar, Inc. is not itself an NRSRO nor does it issue a credit rating on the fund. An NRSRO rating on a fixed-income security can change from time-to-time.

For credit quality, Morningstar combines the credit rating information provided by the fund companies with an average default rate calculation to come up with a weighted-average credit quality. The weighted-average credit quality is currently a letter that roughly corresponds to the scale used by a leading NRSRO. Bond funds are assigned a style box placement of "low", "medium", or "high" based on their average credit quality. Funds with a low credit quality are those whose weighted-average credit quality is determined to be less than "BBB-"; medium are those less than "AA-", but greater or equal to "BBB-"; and high are those with a weighted-average credit quality of "AA-" or higher. When classifying a bond portfolio, Morningstar first maps the NRSRO credit ratings of the underlying holdings to their respective default rates (as determined by Morningstar's analysis of actual historical default rates). Morningstar then averages these default rates to determine the average default rate for the entire bond fund. Finally, Morningstar maps this average default rate to its corresponding credit rating along a convex curve.

For interest-rate sensitivity, Morningstar obtains from fund companies the average effective duration. Generally, Morningstar classifies a fixed-income fund's interest-rate sensitivity based on the effective duration of the Morningstar Core Bond Index (MCBI), which is currently three years. The classification of Limited will be assigned to those funds whose average effective duration is between 25% to 75% of MCBI's average effective duration; funds whose average effective duration is between 75% to 125% of the MCBI will be classified as Moderate; and those that are at 125% or greater of the average effective duration of the MCBI will be classified as Extensive.

For municipal bond funds, Morningstar also obtains from fund companies the average effective duration. In these cases static breakpoints are utilized. These breakpoints are as follows: (i) Limited: 4.5 years or less; (ii) Moderate: more than 4.5 years but less than 7 years; and (iii) Extensive: more than 7 years. In addition, for non-US taxable and non-US domiciled fixed income funds static duration breakpoints are used: (i) Limited: less than or equal to 3.5 years; (ii) Moderate: greater than 3.5 and less than equal to 6 years; (iii) Extensive: greater than 6 years.

Risk and Return
Standard deviation is a statistical measure of the volatility of a portfolio's returns around its mean.

Mean represents the annualized geometric return for the period shown.

Sharpe ratio uses a portfolio's standard deviation and total return to determine reward per unit of risk.

Alpha measures the difference between a portfolio's actual returns and its expected performance, given its beta and the actual returns of the benchmark index. Alpha is often seen as a measurement of the value added or subtracted by a portfolio's manager.

Beta is a measure of the degree of change in value one can expect in a portfolio given a change in value in a benchmark index. A portfolio with a beta greater than one is generally more volatile than its benchmark index, and a portfolio with a beta of less than one is generally less volatile than its benchmark index.

R-squared reflects the percentage of a portfolio's movements that is explained by movements in its benchmark index, showing the degree of correlation between the portfolio and a benchmark. This figure is also helpful in assessing how likely it is that alpha and beta are statistically significant.

Fundamental Analysis
The below referenced data elements are a weighted average of the equity holdings in the portfolio.

The median market capitalization of a subaccount's equity portfolio gives you a measure of the size of the companies in which the subaccount invests.

The Price/Cash Flow ratio is a weighted average of the price/cash-flow ratios of the stocks in a subaccounts portfolio. Price/cash-flow shows the ability of a business to generate cash and acts as a gauge of liquidity and solvency.

The Price/Book ratio is a weighted average of the price/book ratios of all the stocks in the underlying fund's portfolio. The P/B ratio of a company is calculated by dividing the market price of its stock by the company's per-share book value. Stocks with negative book values are excluded from this calculation.

The Price/Earnings ratio is calculated by dividing the market value of the equity assets by the trailing 12 month earnings. The 12 month earnings value comes from multiplying the number of shares and the adjusted trailing 12 months' earnings per share for each equity asset and summing the results.

The Price/Sales ratio is a weighted average of the price/sales ratios of the stocks in the underlying fund's portfolio. The P/S ratio of a stock is calculated by dividing the current price of the stock by its trailing 12 months' revenues per share. In computing the average, Morningstar weights each portfolio holding by the percentage of equity assets it represents.

The return on assets (ROA) is the percentage a company earns on its assets in a given year. The calculation is net income divided by end-of-year total assets, multiplied by 100.

The Return on Equity (ROE) is the percentage a company earns on its shareholders' equity in a given year. The calculation is net income divided by end-of-year net worth, multiplied by 100.

Market Maturity shows the percentage of a holding's common stocks that are domiciled in developed and emerging markets.

The data elements listed below are a weighted average of the fixed income holdings in the portfolio.

Average maturity is used for holdings in the taxable fixed-income category. This is a weighted average of all the maturities of the bonds in a portfolio, computed by weighting each maturity date by the market value of the security.

Credit quality breakdowns are shown for corporate-bond holdings and depict the quality of bonds in the underlying portfolio. The report shows the percentage of fixed-income securities that fall within each credit quality rating as assigned by an NRSRO. Bonds not rated by an NRSRO are included in the not rated (NR) category.

Debt as a percentage of capital is calculated by dividing long-term debt by total capitalization (the sum of common equity plus preferred equity plus long-term debt). This figure is not provided for financial companies.

Duration is a time measure of a bonds interest-rate sensitivity.

Net Margin is a measure of profitability. It is equal to annual net income divided by revenues from the same period for the past five fiscal years, multiplied by 100.

Type Weightings divide the stocks in a given holding's portfolio into eight type designations each of which defines a broad category of investment characteristics. Not all stocks in a given holding's portfolio are assigned a type. These stocks are grouped under NA.

The data elements listed below are a weighted average of the total holdings in the portfolio.

The average expense ratio is the percentage of assets deducted each year for operating expenses, management fees, and all other asset-based costs incurred by the fund, excluding brokerage fees. Please note for mutual funds, variable annuities/life, ETF and closed-end funds we use the gross prospectus ratio as provided in the prospectus. Separate accounts and stocks are excluded from the average expense ratio.

Potential capital gains exposure is the percentage of a holdings total assets that represent capital appreciation.

Investment Risks

International/Emerging Market Equities: Investing in international securities involve special additional risks. These risks include, but are not limited to, currency risk, political risk, and risk associated with varying accounting standards. Investing in emerging markets may accentuate these risks.

Sector Strategies: Portfolios that invest exclusively in one sector or industry involve additional risks. The lack of industry diversification subjects the investor to increased industry-specific risks.

Non-Diversified Strategies: Portfolios that invest a significant percentage of assets in a single issuer involve additional risks, including share price fluctuations, because of the increased concentration of investments.

Small Cap Equities: Portfolios that invest in stocks of small companies involve additional risks. Smaller companies typically have a higher risk of failure, and are not as well established as larger blue-chip companies. Historically, smaller-company stocks have experienced a greater degree of market volatility that the overall market average.

Mid Cap Equities: Portfolios that invest in companies with market capitalization below $10 billion involve additional risks. The securities of these companies may be more volatile and less liquid than the securities of larger companies.

High-Yield Bonds: Portfolios that invest in lower-rated debt securities (commonly referred as junk bonds) involve additional risks because of the lower credit quality of the securities in the portfolio. The investor should be aware of the possible higher level of volatility, and increased risk of default.

Tax-Free Municipal Bonds: The investor should note that the income from tax-free municipal bond funds may be subject to state and local taxation and the Alternative Minimum Tax.

Bonds: Bonds are subject to interest rate risk. As the prevailing level of bond interest rates rise, the value of bonds already held in a portfolio decline. Portfolios that hold bonds are subject to declines and increases in value due to general changes in interest rates.

HOLDRs: The investor should note that these are narrow industry-focused products that, if the industry is hit by hard times, will lack diversification and possible loss of investment would be likely. These securities can trade at a discount to market price, ownership is of a fractional share interest, the underlying investments may not be representative of the particular industry, the HOLDR might be delisted from the AMEX if the number of underlying companies drops below nine, and the investor may experience trading halts.

Hedge Funds: The investor should note that hedge fund investing involves specialized risks that are dependent upon the type of strategies undertaken by the manager. This can include distressed or event-driven strategies, long/short strategies, using arbitrage (exploiting price inefficiencies), international investing, and use of leverage, options and/or derivatives. Although the goal of hedge fund managers may be to reduce volatility and produce positive absolute return under a variety of market conditions, hedge funds may involve a high degree of risk and are suitable only for investors of substantial financial means who could bear the entire loss of their investment.

Bank Loan/Senior Debt: Bank loans and senior loans are impacted by the risks associated with fixed income in general, including interest rate risk and default risk. They are often non-investment grade; therefore, the risk of default is high. These securities are also relatively illiquid. Managed products that invest in bank loans/senior debt are often highly leveraged, producing a high risk of return volatility.

Short Positions: When a short position moves in an unfavorable way, the losses are theoretically unlimited. The broker may demand more collateral and a manager might have to close out a short position at an inopportune time to limit further losses.

Long-Short: Due to the strategies used by long-short funds, which may include but are not limited to leverage, short selling, short-term trading, and investing in derivatives, these funds may have greater risk, volatility, and expenses than those focusing on traditional investment strategies.

Liquidity Risk: Closed-end fund, ETF, and HOLDR trading may be halted due to market conditions, impacting an investor's ability to sell a fund.

Market Price Risk: The market price of ETFs, HOLDRs, and closed-end funds traded on the secondary market is subject to the forces of supply and demand and thus independent of the NAV. This can result in the market price trading at a premium or discount to the NAV which will affect an investor's value.

Market Risk: The market prices of ETF's and HOLDRs can fluctuate as a result of several factors, such as security-specific factors or general investor sentiment. Therefore, investors should be aware of the prospect of market fluctuations and the impact it may have on the market price.

Target-Date Funds: Target-date funds typically invest in other mutual funds and are designed for investors who are planning to retire during the target date year. The fund's target date is the approximate date of when investors expect to begin withdrawing their money. Target-date fund's investment objective/strategy typically becomes more conservative over time primarily by reducing its allocation to equity mutual funds and increasing its allocations in fixed-income mutual funds. An investor's principal value in a target-date fund is not guaranteed at anytime, including at the fund's target date.

High double- and triple-digit returns were the result of extremely favorable market conditions, which may not continue to be the case. High returns for short time periods must not be a major factor when making investment decisions.

Benchmark Disclosure

BarCap US Agg Bond TR USD

This index is composed of the BarCap Government/Credit Index, the Mortgage-Backed Securities Index, and the Asset-Backed Securities Index. The returns we publish for the index are total returns, which include reinvestment of dividends.

MSCI EAFE NR USD

This Europe, Australasia, and Far East index is a market-capitalization-weighted index of 21 non-U.S., industrialized country indexes.

S&P 500 TR

A market capitalization-weighted index of 500 widely held stocks often used as a proxy for the stock market. TR (Total Return) indexes include daily reinvestment of dividends.

USTREAS T-Bill Auction Ave 3 Mon

Three-month T-bills are government-backed short-term investments considered to be risk-free and as good as cash because the maturity is only three months. Morningstar collects yields on the T-bill on a weekly basis from the Wall Street Journal.

42

Fordham Law Review

Volume 74 | Issue 6 Article 4

2006

The Small Laws: Eliot Spitzer and the Way to Insurance Market Reform

Sean M. Fitzpatrick

THE SMALL LAWS: ELIOT SPITZER AND THE WAY TO INSURANCE MARKET REFORM

*Sean M. Fitzpatrick**

"For when you break the great laws, you do not get liberty; you do not even get anarchy. You get the small laws."[1]

INTRODUCTION

In 2004, the property-casualty insurance industry was roiled by a scandal unparalleled in its history, with the world's largest insurance broker, Marsh & McLennan, accused of defrauding customers by "rigging bids" to maximize its own profits. New York State Attorney General Eliot Spitzer's suit against Marsh in October 2004 was the first salvo in a continuing challenge to long-established insurance market practices,[2] and has set in motion a process of regulatory scrutiny and proposed legal reform whose ends are impossible to predict. Indeed, not since another ambitious New York governor-to-be, Charles Evans Hughes, cut a swath through the life insurance industry as chief counsel to the renowned Armstrong Committee—almost exactly a century ago—has the insurance community faced so fundamental a challenge to its structure and ethics.[3] In a curious historical twist, that earlier scandal had its genesis in a grand party thrown

---

* Lecturer in Law, University of Connecticut School of Law; Senior Vice President and Special Counsel, The Chubb Corporation. In his capacity as Special Counsel at Chubb, Mr. Fitzpatrick has participated in Chubb's response to Attorney General Eliot Spitzer's investigation of insurance industry practices and ensuing investigations by other state and federal regulators. Prior to assuming his current position, Mr. Fitzpatrick was chief underwriting officer of Chubb's specialty division from 1999 through 2002. In 2005, Mr. Fitzpatrick also served as president of the Professional Liability Underwriting Society ("PLUS"), the leading trade group in the executive and professional liability insurance industry. The author is indebted to John Degnan, Tom Baker, David Robinson, and Dino Robusto for their helpful comments on early drafts of this Article; the opinions expressed are the author's alone.

1. G.K. Chesterton, Charles Dickens 197 (Schocken Books 1965) (1906).

2. *See* Complaint, 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]

3. *See* Mark J. Roe, *Foundations of Corporate Finance: The 1906 Pacification of the Insurance Industry*, 93 Colum. L. Rev. 639, 656-74 (1993); Adam Winkler, *"Other People's Money": Corporations, Agency Costs, and Campaign Finance Law*, 92 Geo. L.J. 871, 887-91 (2004). More recent insurance scandals, such as the Lloyd's of London "spiral" in the 1980s, pale in significance beside the challenge posed to the industry by Spitzer's investigation. *See* Adam Raphael, Ultimate Risk: The Inside Story of the Lloyd's Catastrophe (1995).

in 1905 by the young heir to control of the Equitable Life Assurance Society, on the site of whose then headquarters—120 Broadway in New York City—now stands the office of Attorney General Spitzer.[4] Regulatory firestorms such as Mr. Spitzer has unleashed are like Pandora's Box: Once opened, even their authors are powerless to control their ultimate effects. And those effects are likely to surprise even the most well-intentioned regulator. In this particular case, a strong argument can be made that an industry-wide ban on the contingent compensation structures targeted by Spitzer would bankrupt hundreds of small insurance agencies in communities throughout America, and lead to the further consolidation of insurance brokerage business in large global firms like Marsh. Insurance is an old-fashioned, surprisingly risk-averse industry. Its fundamental practices have changed little for generations. Indeed, the state of affairs that first drew Attorney General Spitzer's attention—large commercial insurance brokers allegedly manipulating the market for their own benefit—was the result of nothing so much as the coupling of time-honored sales incentive practices developed on Main Street, U.S.A. with an unprecedented level of market power attained by a few global megafirms following a consolidation spree in the 1990s.[5] But it would be an expensive mistake to jump from that observation to the conclusion that all contingent compensation of all insurance producers is necessarily harmful to insurance consumers.

Part I of this Article traces the history of Attorney General Spitzer's investigation into insurance industry practices (the "Spitzer Investigation"). Part II offers an overview of the property-casualty insurance market as it has evolved in the United States and discusses the pros and cons of traditional insurance producer compensation practices.[6] Part III describes regulatory and legislative responses to the Spitzer Investigation, largely focused in interstate organizations such as the National Association of Insurance Commissioners and the National Conference of Insurance

---

4. *See* Patricia Beard, After the Ball: Gilded Age Secrets, Boardroom Betrayals, and the Party that Ignited the Great Wall Street Scandal of 2005, at 14 (Perennial ed. 2004).

5. Among the many ironies of the current investigations—which include a substantial focus on threats to competition in the insurance market—is the fact that the consolidation of more than sixty percent of the global insurance brokerage market for large corporate risks into three megafirms occurred in the 1990s without so much as a peep from federal or state antitrust authorities. *See Oversight Hearing on Insurance Brokerage Practices, Including Potential Conflicts of Interest and the Adequacy of the Current Regulatory Framework: Hearing Before the Subcomm. on Financial Management, the Budget, and International Security, Comm. on Governmental Affairs*, 108th Cong. 55 (2004) (statement of Eliot Spitzer, State of New York Attorney General), *available at* http://www.oag.state.ny.us/press/statements/insurance_investigation_testimony.pdf [hereinafter U.S. Senate Testimony].

6. Although Attorney General Spitzer and other regulators have also launched investigations into practices in the market for health and employee benefits insurance, *see, e.g.*, Press Release, Office of N.Y. State Attorney Gen. Eliot Spitzer, Life, Disability Broker Charged with Fraud, Antitrust Violations (Nov. 12, 2004), *available at* http://www.oag.state.ny.us./press/2004/nov/nov12a_04.html, this Article will focus on the property-casualty insurance market.

Legislators. Finally, Part IV sets forth the author's suggestions for a simple, voluntary reform that would increase transparency for insurance consumers while avoiding the pitfalls likely to attend more draconian solutions.

I. THE SPITZER INVESTIGATION

The Spitzer Investigation began without fanfare on February 10, 2004, when the Washington Legal Foundation ("WLF"), a free-market-oriented advocacy organization, sent a letter to the state insurance commissioners and attorneys general of New York and California. In its letter, the WLF voiced concerns with "two potentially damaging practices engaged in by some in the insurance brokerage industry":[7] namely, the use of "placement service agreements" by brokers to obtain additional compensation from insurers based on the volume of business placed with them and the alleged "leveraging" by brokers of primary insurance production "to procure an insurance company's highly lucrative reinsurance [placement] business."[8] While the WLF letter focused on so-called placement service agreements ("PSAs")—a usage coined by Marsh and not widely used elsewhere in the insurance industry—the thrust of the WLF's complaint was that so-called "contingent commissions" paid by insurers to brokers based on the brokers' achievement of premium volume and profitability goals "can compromise the broker's fiduciary duty to represent the best interests of their clients, and create incentives for brokers to refer business to companies that will make them more money."[9] Similarly, the WLF argued that a broker's leveraging of its ability to refer primary insurance business to obtain reinsurance brokerage engagements from insurers could lead to a similar conflict of interest.[10]

In their typical form, "contingent commissions" are payments made by an insurer to an insurance agency or brokerage for success in achieving stipulated levels of premium volume and profitability on its overall book of business with that carrier.[11] Such payments are in addition to the "standard commissions"—typically ranging from ten to twenty percent of the total

7. Letter from Daniel J. Popeo, Chairman and Gen. Counsel, Washington Legal Found., to Gregory V. Serio, N.Y. State Superintendent of Ins. (Feb. 10, 2004) (on file with author) [hereinafter WLF Letter]. The Washington Legal Foundation ("WLF") letter credited a December 2003 article by John Dizard of the Financial Times as prompting its interest in these issues, *see* John Dizard, *With Brokers like These Who Needs Enemies?*, Fin. Times, Dec. 15, 2003, at 27, although it also made note of a January 2004 J.P. Morgan Securities analyst report that focused on the significance of contingent commissions in the overall compensation of large brokers, *see* Hugh Warns et al., J.P. Morgan Sec., Insurance—Non-life: Contingents May Be Smaller, But More Prominent in 2004 (2004), *available at* http://www.cdfe.org/jpm%20cont%20comm%201041.pdf.

8. WLF Letter, *supra* note 7, at 2.

9. *Id.* at 1.

10. *Id.* at 2-3.

11. *See* J. David Cummins & Neil A. Doherty, The Economics of Insurance Intermediaries 2 (2005), *available at* http://www.insurancejournal.com/downloads/WhartonStudy_2005.05.20.pdf.

policy premium—that are paid by the carrier on each discrete insurance transaction. Marsh's innovation, developed in the late 1990s, was a "Placement Service Agreement" that ostensibly compensated the broker based on services provided to the insurer and was calculated based on premium volume alone, without regard to the ultimate profitability of the business produced by Marsh. A few other large brokers adopted the PSA terminology, sometimes as an acronym for "Profit Sharing Agreement," but profitability remained a component of most non-Marsh PSAs.[12]

Interestingly, in view of what followed, the WLF's concern was limited to conflicts involving insurance brokers, who "are paid to advocate for their customers, not themselves, and certainly not for the insurance companies with whom they place their business."[13] The WLF did not address the more complicated issue of independent insurance agents, intermediaries who are contractually bound to represent the interests of insurance carriers that appoint them.[14] As I discuss in the next section, intermediaries operating solely as true "brokers" are relatively rare in the United States, and tend to be centered in certain geographic areas, most notably New York City and San Francisco, whose insurance markets developed from maritime roots and not surprisingly borrowed their structure from the original "broker market," Lloyd's of London. Elsewhere in the United States, the demands of distance led to the development of an "agency system" whereby intermediaries were empowered to act on behalf of one or more insurers in performing functions not permitted to true brokers, such as binding and issuing policies.[15] Independent insurance agents continue to place the bulk of property-casualty policies in this country.[16]

---

12. In 2004, as questions about contingent commissions arose, Marsh and its largest competitor, Aon Corporation, briefly adopted the term "Market Service Agreement" or "MSA" to describe their contingent commission agreements with carriers, but this nomenclature was never adopted by the industry at large.

13. WLF Letter, *supra* note 7, at 3. In fact, contingent compensation paid to brokers had been subjected to regulatory scrutiny by New York State as recently as 1998, when the state's Department of Insurance issued a Circular Letter requiring brokers to disclose contingent commissions to their clients. *See* N.Y.S. Dept. of Ins., Circular Letter No. 22 (Aug. 25, 1998).

14. *See* 7 Eric Mills Holmes, Holmes' Appleman on Insurance 2d: Law of Insurance Agents § 47.5, 326 (1998) ("Bluntly stated, an 'insurance agent' represents the insurance company, whereas an 'insurance broker' represents the insured, although the question whether one is an insurance agent or broker is a question dependent on the particular facts." (footnotes omitted)); Int'l Risk Mgmt. Inst., Professional Liability Insurance, at XV.C.1 (2003) [hereinafter IRMI] ("Traditionally, the difference between insurance agents and brokers is that agents are considered representatives of insurance companies while brokers are thought of legally as representatives of insureds."); Sean M. Fitzpatrick, *Fear Is the Key: A Behavioral Guide to Underwriting Cycles*, 10 Conn. Ins. L.J. 255, 269 n.47 (2004); *see also* Krumme v. Mercury Ins. Co., 20 Cal. Rptr. 3d 485, 488-90 (Cal. Ct. App. 2004) (discussing the broker-agent distinction under California law). For purposes of this Article, the terms "producer" and "intermediary" will be used to describe both brokers and agents.

15. *See generally* Holmes, *supra* note 14, §§ 44.1-.7.

16. *See, e.g.*, Cummins & Doherty, *supra* note 11, at 8; IRMI, *supra* note 14, at XV.C.5; *see also* Owens v. Aetna Life & Cas. Co., 654 F.2d 218, 240 (3d Cir. 1981).

Spitzer's interest in potential abuses in the insurance industry was further whetted by an anonymous letter received by his office on March 30, 2004. The letter, postmarked from a New York suburb and signed only "Concerned," made damning allegations about Marsh's PSAs, asserting, "The point is to appear as if Marsh is providing a service to the insurance market rather than the reality which is that Marsh is receiving major income for directing business to preferred providers/insurance markets."[17] Spitzer's office reportedly issued a subpoena to Marsh within three days of receiving this corroboration of the WLF's allegations.[18]

The Spitzer Investigation became public on April 22, 2004, when Aon, the world's second-largest insurance broker, reported that it had received a subpoena from Spitzer's office inquiring about its acceptance of contingent commissions from insurers.[19] Soon thereafter, Marsh and Willis Group Holdings, the world's third largest broker, reported receiving similar subpoenas.[20] Within a month, Spitzer's office had broadened its probe to include insurance companies as well as brokers, issuing a number of subpoenas to large property-casualty insurers.[21]

The New York Attorney General broadened his investigation into two additional areas in mid-2004. In a second round of subpoenas issued to insurance carriers in August 2004, the Attorney General's office first sought information regarding insurers' compensation arrangements with independent insurance agents, as opposed to brokers.[22] At the same time, documents were requested concerning the second alleged source of insurance market distortion raised by the WLF—the "tying" by an insurance intermediary of retail business production to an insurance

17. *See* Steve Fishman, *Inside Eliot's Army*, N.Y. Mag., Jan. 10, 2005, *available at* http://www.newyorkmetro.com/nymetro/news/politics/newyork/features/10815/index.html; *see also* Peter Elkind, *Spitzer's Crusade*, Fortune, Nov. 15, 2004, at 130.

18. Fishman, *supra* note 17.

19. Press Release, Aon Corp., State of New York Seeks Information Regarding Compensation Agreements Between Insurance Brokers and Insurance Companies (Apr. 22, 2004), *available at* http://www.aon.com/about/news/press_release/pr_000100C7.jsp.

20. David K. Bradford, Advisen, Placement Service Agreements: Big Brokers Under Fire 1 (2004); *see also* Gretchen Morgenson, *Hat Trick: A 3rd Unit of Marsh Under Fire*, N.Y. Times, May 2, 2004, § 3, at 9. At least one smaller New York City broker, Kaye Associates (a subsidiary of Hub International), was also subpoenaed in this first wave. *Id.* California Insurance Commissioner John Garamendi was the first regulator outside New York State to follow Spitzer's lead in announcing his own investigation. Joseph B. Treaster, *An Inquiry into Insurance Payments and Conflicts*, N.Y. Times, Apr. 28, 2004, at C1. Others would follow; at this writing more than twenty other states have initiated investigations of insurance market conduct.

21. Joseph B. Treaster, *Inquiry Widens; Insurance Brokers May Face Change*, N.Y. Times, May 18, 2004, at C1. The author's employer, Chubb Corporation, was among the insurance companies subpoenaed by Attorney General Spitzer's office. *See* Press Release, The Chubb Corp., Chubb to Comply with Subpoena Regarding Broker Compensation (May 17, 2004), *available at* http://www.chubb.com/marketing/chubb1824.html.

22. A sample subpoena is on file with the author.

carrier's willingness to retain that intermediary's reinsurance brokerage affiliate to place its reinsurance.[23]

Perhaps the most fateful day in determining the ultimate course of the Spitzer Investigation was September 9, 2004, when a New York University law student working as an intern in Spitzer's office unearthed an e-mail indicating that Marsh had conspired with insurance carriers to rig bids on insurance renewals: obtaining false, inflated quotes from complicit markets to enable Marsh to direct business to a chosen market with an eye toward maximizing its PSA payments.[24] The quid pro quo for such cooperation from the "losing" markets would be protection by Marsh on its own renewals through similar means. This evidence of outright fraud and market manipulation provided the catalyst that enabled Spitzer's staff to crystallize a more unfocused discomfort with insurance market sales incentives (including contingent commissions, loans to producers, funding of producers' sales staffs, leveraging reinsurance broking engagements, etc.) into a broader challenge of disclosure practices in the property and casualty ("P&C") market.

This challenge manifested on October 14, 2004, when Spitzer filed suit against Marsh.[25] The thrust of Spitzer's complaint was two-fold. First, he alleged that Marsh's practice of accepting contingent commissions led the broker to elevate its own interests above those of its clients, by "steering" business to carriers based more on the economics of Marsh's PSAs than on the needs of the customer.[26] Second, Spitzer alleged several instances of outright bid rigging by Marsh in the Excess Casualty insurance market,

23. *See* WLF Letter, *supra* note 7, at 2-3. The Attorney General issued a more comprehensive round of subpoenas focused on reinsurance tying allegations in October 2004. Whereas Spitzer's investigation of alleged abuses related to placement service agreements ("PSAs") focused largely on activities at Marsh, this facet of his investigation appeared to focus primarily on business practices at Aon. *See* Timothy L. O'Brien & Joseph B. Treaster, *Spitzer Goes Hunting for His Next Trophy*, N.Y. Times, Oct. 31, 2004, § 3, at 1.

24. Kate Kelly, *In Spitzer's Office, Hours of Drudgery, Moments of 'Gotcha!'*, Wall St. J., Oct. 27, 2004, at A1.

25. *See* Marsh Complaint, *supra* note 2.

26. *Id.* ¶¶ 7-8, 14-42. In an amusing footnote to this story, Spitzer was embarrassed in April 2005 when Reuters reported that his gubernatorial campaign had paid the Google search engine firm to direct users searching the term "AIG"—an insurance company then under intensive investigation by Spitzer's office—to a link to the "Spitzer for NY Governor" campaign website, which featured the message: "Good Guys Can Finish First. Sign up Now to Join the Fight!" *NY's Spitzer Gubernatorial Campaign Goes to Google*, Reuters, Apr. 6, 2005 (on file with author, together with screen capture of link as it appeared). Within hours of the Reuters report, the link had been removed and Spitzer's spokesman had publicly disavowed the "relatively low-level campaign staffer" responsible. *See Spitzer Pulls Campaign Ad Off Google AIG Link*, Reuters, Apr. 6, 2005 (on file with author). To his credit, Attorney General Spitzer quickly renounced this campaign device, no doubt appreciating the irony that his hapless staffer's Google gambit had in effect caused his campaign to make an undisclosed payment to an intermediary for the purpose of steering consumers of information services to its website.

involving collusion with a number of major insurers.[27] On the same day he filed his complaint against Marsh, Spitzer underscored the seriousness of his allegations by announcing guilty pleas by two employees of insurer American International Group ("AIG"), who admitted to participating in criminal bid rigging.[28]

It would eventually become clear that, while Spitzer had discovered a serious problem in one niche of the market, outright bid-rigging behavior appears to have been limited to the excess casualty sector, where the combination of a commoditized product and a highly concentrated brokerage market created conditions ripe for abuse. But, even if we accept that the bid rigging admittedly conspired in by a number of Marsh employees (as well as underwriters at several major insurance companies) was an aberration, how do we account for it? Why would executives at Marsh risk the professional reputation of the firm, as well as the loss of substantial clients and the long-term income stream they represented, in return for the transient benefit of achieving any one year's PSA payment from a particular insurer? From Marsh's standpoint, the answer lies in the unfortunate way the broker apparently structured its profit centers, and the perverse incentives that flowed from that structure. By setting up its Global Broking division so that its sole source of income was PSA payments from carriers, and by assigning internal "credit" for the normal income derived from standard commissions to a separate Client Advisory division, Marsh all but guaranteed that its employees in Global Broking would elevate the maximization of PSA payments over all other considerations, including the long-term reputation and even financial well-being of the firm. As I discussed at length in an earlier article on the insurance underwriting cycle, insurance professionals respond to the incentives provided them, and many of the "problems" the insurance industry contends with—from "boom" and "bust" market cycles to the recent bid-rigging scandals—can be traced fairly directly to ill-conceived compensation structures.[29]

Other commentators, notably Professor John Coffee, have also observed that even a firm whose very value derives from the trust of the marketplace in its integrity can unwittingly place its franchise at risk by creating a bureaucratic culture where profit centers have powerful short-term economic interests that are inconsistent with the firm's long-term interests,

---

27. Marsh Complaint, *supra* note 2, ¶¶ 43-66. Excess Casualty is a type of multi-peril liability insurance that sits in excess of more specific primary liability policies; it is a highly commoditized product available from many insurance carriers.

28. Press Release, Office of N.Y. State Attorney Gen. Eliot Spitzer, Investigation Reveals Widespread Corruption in Insurance Industry (Oct. 14, 2004), *available at* http://www.oag.state.ny.us/press/2004/oct/oct14a_04.html; Matthew Goldstein, *Spitzer Charges Marsh & McLennan in Insurance Racket*, The Street.com, Oct. 15, 2004, http://www.thestreet.com/markets/matthewgoldstein/10187969.html. Additional guilty pleas followed in the ensuing months. At this writing, a total of seventeen individuals had pled guilty to participation in bid rigging, and eight more were under indictment. Press Release, Office of N.Y. Attorney Gen. Eliot Spitzer, Insurance Execs Indicted for Bid-Rigging, Fraud (Sept. 15, 2005), *available at* http://www.oag.state.ny.us/press/2005/sep/sep15a_05.html.

29. *See generally* Fitzpatrick, *supra* note 14.

and where the necessary internal and external controls to manage such an institutional tension are lacking.[30] Coffee also astutely points out that this risk is exacerbated where, as in the large broker market, a limited number of competitors reduces the risk that questionable practices will be "outed" by marketplace rivals.[31] Marsh should, perhaps, have seen warning signs—friction between its Global Broking and Client Advisory arms was an open secret in the P&C industry—but beyond reportedly trying to adjust the incentives of its client advisors to gain their buy-in to the Global Broking business model,[32] Marsh plainly did not do enough to safeguard against the risks of that model. Indeed, given the incentives provided and the absence of necessary countermeasures, it speaks well of Marsh's employees outside the Excess Casualty area that they apparently did not succumb to the temptation to game Marsh's compensation system at the expense of its clients.

But what of the underwriters who conspired with employees of Marsh Global Broking in rigging bids? Were they so driven to meet premium production goals that they became blind to the fundamental wrongness of their conduct? While this is possible, I suspect the real answer is more subtle, and lies in the way that ethical and cultural norms are communicated within the insurance industry. Insurance is an industry without an established, written set of rules governing business practices (unlike, for example, law or accounting). New initiates to the insurance business typically receive their training in an informal way, from more senior employees, and there is a tendency in the industry to accept business practices of long standing at face value, without examining their ethical implications. The structure of the insurance market—with intermediaries acting as the gatekeepers of new business opportunities—also creates some confusion among underwriters as to whom precisely their "customer" is: the producer they must please in order to sell policies or the end user of those policies. Taken together, these influences make it sadly likely that at least some of the underwriters who have pled guilty to providing fictitious quotes to Marsh never thought of their behavior as unethical, let alone criminal. Rather, they probably provided such quotes when asked after being told "this is just how it's done," well aware, as they no doubt were, of the importance of maintaining good relations with the world's largest insurance producer. Again, the surprising thing may be not that bid rigging occurred among a small number of players in the excess casualty insurance

30. *See* John C. Coffee, Jr., *Understanding Enron: "It's About the Gatekeepers, Stupid"*, 57 Bus. Law. 1403, 1409-16 (2002). It is worth noting in this context that another potential check on broker misbehavior—regular customer contact with their insurers—was absent to a unique degree in the Excess Casualty market. Excess Casualty is a product involving infrequent claims, and requires neither the regular servicing of traditional insurance products like property insurance, nor the in-depth annual underwriting process (often involving face-to-face meetings with underwriters) typical of complex casualty lines such as directors' and officers' liability insurance.

31. *Id.* at 1414-15.

32. Marsh Complaint, *supra* note 2, ¶ 39.

market, but that similar practices did not develop in other sectors of the P&C market where limited intermediary channels, a commoditized product, and dangerous financial incentives also existed.

In any event, Spitzer's combination of a broad-based attack on contingent commissions, long a familiar and public feature of the insurance market,[33] with an indictment of outright fraud in the form of bid rigging, was neatly done. And, while some commentators pointed out the questionable logic of meting out equal condemnation to the two very different practices,[34] the distinction was largely overlooked in the media barrage that followed Spitzer's filing against Marsh. The term "kickback" was almost uniformly adopted in press accounts of the Marsh Complaint and even in regulatory pronouncements from other states,[35] notwithstanding the New York Attorney General's own initial reluctance to use the term.[36]

Throughout late 2004 and early 2005, however, Spitzer sent conflicting signals as to how he viewed contingent commissions in and of themselves, untainted by association with bid rigging or other types of fraud. One could infer from the Marsh Complaint that Spitzer believed such payments were inevitably corrupting of the insurance market. At other times, however, the Attorney General has expressly stated that his concern is not with contingent commissions per se, but rather with inadequate disclosure to customers of such payments. In a speech to the National Press Club on January 31, 2005, for example, Spitzer acknowledged that, while the way contingent commissions were used by mega-brokers such as Marsh was wrong, "in other parts of the industry, they may be appropriate."[37] "There

33. Indeed, some of the most vocal critics of contingent commissions as "kickbacks" have overlooked the fact that their evidence was provided by the insurers themselves, who publicly report on contingent commissions paid in their annual statutory statements filed with each state's insurance department. *See, e.g.*, J. Robert Hunter, Consumer Fed'n of Am., Contingent Insurance Commissions: Implications for Consumers 2 (2005), *available at* http://64.233.179.104/search?q=cache:ihRWPXhE5t0J:www.consumerfed.org/.

34. *See, e.g.*, *Eliot's Insurance Policy*, Wall St. J., Oct. 21, 2004, at A18; Henry G. Manne, *Regulation 'In Terrorem,'* Wall St. J., Nov. 22, 2004, at A14.

35. *See, e.g.*, Theo Francis, *Spitzer Charges Bid Rigging in Insurance*, Wall St. J., Oct. 15, 2004, at A1; *see also* U.S. Senate Testimony, *supra* note 5, at 2; Press Release, Cal. Dep't of Ins., Insurance Commissioner John Garamendi Sues Broker and 4 Major Insurers over Secret Commissions and Kickback Schemes that Netted "Millions of Dollars" (Nov. 28, 2004), *available at* http://www.insurance.ca.gov/0400-news/0100-press-releases/0090-2004/release095-04.cfm.

36. Earlier in his investigation, Attorney General Spitzer had been reluctant to use the word "kickback" in describing contingent commissions. *See* Interview with Eliot Spitzer (WXXI-Rochester, NY radio broadcast May 28, 2004) (transcript available at http://www.wxxi.org/ntk/Transcripts/2004/0528.html). But, despite eschewing the term in the Marsh Complaint itself, Spitzer had adopted the media's popular term by the time of his November 2004 U.S. Senate testimony. Of course, as recounted above, his office had discovered Marsh's bid rigging in the interim.

37. Press Release, Nat'l Assoc. of Prof'l Ins. Agents, PIA National Encouraged by Remarks Made by NY Attorney General Eliot Spitzer on Contingent Commissions (Jan. 31, 2005), *available at* http://www.pianet.com/NewsCenter/PressReleases/1-31-05.htm. A recording of Spitzer's entire January 31, 2005, National Press Club speech is available at http://www.npr.org/templates/story/story.php?storyId=4472927 (last visited Apr. 17, 2006).

might be other contexts," he added, "where contingent agreements might not usurp decision-making and fiduciary duty."[38]

In any event, Spitzer's filing of the Marsh Complaint set off a nationwide wave of investigations, and regulatory and legislative initiatives to reform insurance brokerage practices.[39] Spitzer's subsequent $850 million settlement with Marsh, announced in January 2005 on the same day as his National Press Club speech, lent credence and momentum to these investigations.[40]

The New York Attorney General retained the initiative even as other regulators joined the fray, issuing new rounds of subpoenas focused on (1) allegations of a conspiracy among lawyers' malpractice insurers to boycott plaintiffs' class action law firms,[41] and (2) "finite risk" and other "nontraditional" insurance products allegedly utilized by some companies to manipulate their financial results.[42] The latter of these prompted the U.S. Securities and Exchange Commission to begin its own investigation, following up on its earlier enforcement actions against insurance industry giant AIG regarding finite risk products it had sold with a broad investigation paralleling Spitzer's. Spitzer's office also broadened its focus from contingent commission arrangements to other financial dealings between insurers and producers, including insurer funding of individual producer's salaries and incentives to limit marketing of renewals. These facets of the Spitzer investigation gained prominence when the Attorney General simultaneously sued Aon and announced a $190 million settlement

38. *Spitzer: All Marsh Funds to Go to Insureds; Not All Insurance People or Contingencies are 'Bad,'* Insurance Journal, Jan. 31, 2005, http://www.insurancejournal.com/news/national/2005/01/31/50488.htm.

39. At this writing, more than twenty states have initiated formal investigations into insurance market conduct. Both the National Conference of Insurance Commissioners ("NAIC") and National Conference of Insurance Legislators ("NCOIL") have promulgated model legislation directed at disclosure to consumers of insurance producer compensation. *See infra* Part III. Spitzer himself extended his investigation of contingent commissions and bid rigging into the health insurance market in November 2004, with a suit against a leading health and benefits broker, Universal Life Resources. *See* Press Release, Office of N.Y. State Attorney Gen. Eliot Spitzer, Life, Disability Broker Charged with Fraud, Antitrust Violations (Nov. 12, 2004), *available at* http://www.oag.state.ny.us/press/2004/nov/nov12a_04.html.

40. *See* Press Release, Office of N.Y. State Attorney Gen. Eliot Spitzer, 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.

41. *See* Press Release, The St. Paul Travelers Cos., St. Paul Travelers Receives Subpoena Relating to Lawyers' Professional Liability Insurance (Dec. 10, 2004), *available at* http://investor.stpaultravelers.com/phoenix.zhtml?c=177842&p=irol-newsArticle&ID=653573&highlight=; Insurers Subpoenaed Over Lawyer Liability Coverage, http://blogs.advisen.com/blojsom/blog/default/LPL%20Controversy%20News/ (Dec. 14, 2004, 09:39 EST).

42. *See* Gretchen Morgenson, *Next Up for Spitzer: Funny Numbers*, N.Y. Times, Nov. 21, 2004, § 3, at 1. These inquiries, while raising fascinating questions in and of themselves, are beyond the scope of this Article.

with the brokerage firm in March 2005.[43] In April 2005, Spitzer reached a similar $50 million settlement with the number three broker, Willis.[44] Marsh, Aon, and Willis each agreed to essentially the same set of "business reforms" as part of their settlements; chief among these was a ban on their acceptance of contingent commissions and other forms of incentive compensation from insurers.[45]

It is not yet clear whether Attorney General Spitzer intends to press this reform on smaller producers, including independent agents, although his remarks at the National Press Club indicate that he may stop short of doing so.[46] If he does hesitate, what potential consequences of such an outright

43. *See* Press Release, Office of N.Y. State Attorney Gen. Eliot Spitzer, 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.

44. *See* Press Release, Office of N.Y. State Attorney Gen. Eliot Spitzer, State Settles Probe of Willis: Third-Largest Insurance Broker to Pay $50 Million and Adopt Reforms (Apr. 8, 2005), *available at* http://www.oag.state.ny.us/press/2005/apr/apr08b_05.html. Aon and Willis had forsworn accepting contingent commissions almost immediately following Spitzer's suit against Marsh. *See* Press Release, Aon Corp., Aon Eliminates Contingent Commissions; CEO Calls for New Approach (Oct. 22, 2004), *available at* http://www.aon.com/about/news/press_release/pr_0079D39E.jsp; 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. Several other large brokerage firms followed suit in the weeks that followed. *See, e.g.*, 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.corporate-ir.net/media_files/irol/10/104111/AJG_Q3_2004_Earnings_Release.pdf.

45. The "business reforms" agreed to by Marsh, Aon, and Willis include the following: written disclosure and client acknowledgement of standard fee or commission arrangements; prohibition of contingent commissions; prohibition of any "pay to play" charges to insurers; prohibition of "bid-rigging" arrangements; prohibition of reinsurance brokerage "leveraging"; prohibition of inappropriate use of wholesalers (i.e., running placements through affiliated Wholesale brokers to generate additional commissions); mandatory disclosure to clients of all insurance quotes procured and related compensation; and implementation of written standards of conduct and ethics training. The full texts of the Marsh, Aon, and Willis settlement agreements are available on Attorney General Spitzer's web site at http://www.oag.state.ny.us/press/agpress05.html (last visited Apr. 17, 2006).

46. David Brown, the chief of Spitzer's investor protection bureau, has taken a more aggressive public stance, noting, "[I]t'll be interesting to see if the smaller operators reform themselves. If they don't, maybe we'll have a more extended inquiry on our hands." Elkind, *supra* note 17, at 134. As recently as October 2005, some in Spitzer's office were remarking that "we're only in the fourth inning . . . [; t]here's been very little resolution and there's no plan to walk away from this." Ellen Kelleher & Andrea Felsted, *A Year on, Spitzer Keeps up the Pressure*, Fin. Times, Oct. 14, 2005, at 28. As this Article went to press, Attorney General Spitzer provided additional insights into his approach to the general issue of contingent commissions when, on February 9, 2006, he announced a $1.6 billion settlement with AIG that resolved, inter alia, allegations of bid rigging and abusive contingent commission practices by that firm. *See* Press Release, Office of N.Y. State Attorney Gen. Eliot Spitzer, AIG Settles Fraud, Bid-Rigging and Improper Accounting Charges (Feb. 9, 2006), *available at* http://www.oag.state.ny.us/press/2006/feb/feb09a_06.html. In the settlement agreement, Spitzer exacted AIG's agreement to support legislation banning contingent compensation outright (an unlikely outcome, as discussed in Part III, *infra*) and,

ban should concern him? Recalling Chesterton's observation, does the fact that some participants in the insurance market broke certain "great laws" (the Eighth and Tenth Commandments come to mind[47]) leave us no recourse but a fifty-state patchwork of new producer compensation and disclosure regulations, "small laws" that may disrupt the fundamental economics of the insurance industry? If this prospect is unappealing, what alternatives are available to curb insurance market abuses that will not threaten lasting damage to a vital sector of the economy?

But before we reach these questions, it will be helpful to describe the U.S. property and casualty insurance market as it had evolved prior to the Spitzer Investigation, and to consider the economics of that market.

II. THE INSURANCE MARKET AND CONTINGENT COMMISSIONS

To trace the origins of the modern U.S. insurance market, we must begin at Edward Lloyd's famous coffee house in seventeenth-century London. Brokers first appeared in this market as a convenient mechanism for communication among the risk takers, or "underwriters," who collectively insured individual ships and cargoes.[48] This system of intermediaries permitted providers of capital to the insurance market to transact business with each other and with purchasers of insurance efficiently, sparing each underwriter the expense of maintaining its own staff of "office keepers." Business at Lloyd's was conducted, much as it is today, within the tight

more interestingly, its agreement to forswear contingent commissions in any line of insurance where the Attorney General could demonstrate based on industry data that the insurance companies collectively writing sixty-five percent or more of the premium volume either did not pay contingents or had agreed to abide by the same commitment as AIG. *See* Agreement Between the Attorney General of the State of New York and American International Group, Inc. and its subsidiaries (collectively, "AIG") dated January 18, 2006, *available at* http://www.oag.state.ny.us/press/2006/feb/signedSettlement.pdf. Attorney General Spitzer will presumably utilize his prosecutorial leverage to encourage competitors of AIG in various insurance lines to agree to similar terms in the coming months, but it remains to be seen whether the sixty-five percent threshold stipulated will be practically attainable in any substantial segment of the widely fragmented property and casualty ("P&C") market. In any event, this mechanism demonstrates both Attorney General Spitzer's view that the insurance market would be a fairer playing field without contingent commissions and his recognition that no single insurer can "unilaterally disarm" in this area. Of course, it also implies that Spitzer realizes he will need to continue to seek reform through investigatory pressure on individual companies, given the low probability of sweeping legislative or regulatory reform in this area.

47. The Eighth Commandment is "Thou shalt not steal"; the Tenth Commandment is "Thou shalt not covet thy neighbor's [goods]." *See Exodus* 20:15, 20:17 (King James). My fellow Roman Catholics will, of course, recognize the former as the Seventh Commandment, but—as our story is about to turn to London—I will defer to our Anglican forebears in this respect.

48. Raphael, *supra* note 3, at 33-36. Amusingly, in light of our current subject, Adam Raphael points out that early brokers at Lloyd's adopted the euphemism "office keepers, broking being regarded as disreputable." *Id.* at 35. Insurance itself, of course, has even older roots, dating from at least as far back as the eighteenth century B.C.E., when an early form of marine insurance known as "bottomry" was enshrined in the Code of Hammurabi in ancient Babylonia. *See* Peter L. Bernstein, Against the Gods: The Remarkable Story of Risk 92-95 (1996).

confines of a small physical marketplace. Initially, underwriters could personally oversee all aspects of their insurance transaction and, as the number of participants in the market grew, "syndicates" of capital providers, or "names," were formed. These syndicates would choose one among their number to serve as the "active underwriter" for the group, and these active underwriters carried on the business for centuries thereafter on much the same personal basis that their forebears had.[49] In this system, brokers were accorded a distinct role within the Lloyd's constitution, as presenters of business opportunities without underwriting authority.[50]

Interestingly, in this same period, the underwriters at Lloyd's employed a distinct, worldwide network of "Lloyd's agents" to provide reports from, and represent their interests in, remote locations.[51] By 1880, there were more than 1000 Lloyd's agents around the world, whose authority extended to the approval of claims payments.[52] The transplantation to and subsequent development of this system in the United States laid the seeds of the most challenging puzzle now facing Attorney General Spitzer.

A. *Brokers and Agents in the U.S. Property-Casualty Market*

The United States' property and casualty insurance market had its birth in marine underwriting, much like its London parent, and the roles and legal relations of the various players in the marketplace were similarly derivative of the London model.[53] Locally based fire insurance companies were also organized on a mutual-ownership basis in several seaboard cities during the colonial period.[54] When the American Revolution removed British restrictions on the formation of joint stock companies, substantial underwriting companies were organized for the first time in the new United States, beginning with two marine insurance companies in Charleston, South Carolina, chartered in 1776. With the formation of the Insurance Company of North America in Philadelphia in 1794, a uniquely American insurance marketplace—characterized by stock and mutual companies dwarfing individual Lloyd's syndicates in size—began in earnest.[55]

At the same time, the United States was expanding at a rapid pace in both geography and population. The large new insurers in cities such as Philadelphia and Hartford needed a mechanism to do business across distances that dwarfed not only the City of London, but all of England. The Lloyd's agency system, modified to the American setting, provided the solution. While insurance intermediaries acting as true brokers—that is,

---

49. *See generally* Raphael, *supra* note 3, at 37-46.

50. *Id.* at 45, 64-65.

51. *Id.* at 27.

52. *Id.* Eventually, Lloyd's extended its delegation of authority to some agents to include underwriting, or "binding," authority. *See, e.g.*, Lloyd's Binding Authority Registration, http://www.lloyds.com/lloyds_market/market_participants/coverholders/binding_authority_registration_BAR (last visited Mar. 2, 2006).

53. Robert J. Gibbons et al., Insurance Perspectives 12-14 (1992).

54. *Id.* at 14-15.

55. *Id.* at 13.

representing insurance purchasers—remained the rule in traditional maritime insurance markets such as New York and San Francisco, elsewhere in the United States a new model predominated, with intermediaries contracting to represent the interests of one or more insurance companies in marketing their products.[56] Ultimately, the two most common types of insurance intermediaries in the United States were "exclusive agents," who were employed by and represented only one insurer, and "independent agents," who represented multiple insurers.[57] Because the latter distribution model required a departure from traditional principles of agency law, particularly in the area of ownership of customer information and renewal rights, the independent agency system developed a formalized set of legal rules known as the American Agency System.[58]

Notwithstanding the historical, geographical, and conceptual distinctions between brokers and independent agents, however, the legal distinction between the two has grown imprecise in recent decades. Indeed, one can hardly locate an in-depth legal analysis of the broker-agent distinction that does not feature words such as "blurry" or "cloudy."[59] Some states no longer even recognize multiple species of insurance intermediaries, denominating all intermediaries simply as "agents" or "producers" (this is the case, for example, in Connecticut); other states, including major insurance markets such as New York and California, continue to recognize two separate species of intermediaries.[60] In a similar vein, some courts and commentators have simply given up, conflating brokers and independent agents for analytical purposes.[61] Even in jurisdictions where brokers and independent agents have largely been viewed as synonymous, however,

56. *See* Deborah S. Freeman & Celia Eggert, *Exploration of Policyholder Information Ownership Rights Under the Three Existing Insurance Agency Systems in the United States*, 23 W. New Eng. L. Rev. 409, 411-16 (2002).

57. *Id.* at 413-19.

58. *Id.* at 415; *see In re* Roy A. Dart Ins. Agency, Inc., 5 B.R. 207, 209-10 (Bankr. D. Mass. 1980); *In re* Estate of Corning, II, 488 N.Y.S.2d 477, 481-82 (App. Div. 1985); *see also* Nationwide Mut. Ins. Co. v. Comm'r of Ins., 491 N.E.2d 1061, 1063 (Mass. 1986) (distinguishing the "Exclusive Agency System" from the "American Agency System").

59. *See, e.g.*, IRMI, *supra* note 14, at XV.C.1 (noting that "the distinctions between the legal duties owed by agents and brokers have, in recent years, become blurred"); Colin Sammon, Comment, *Insurance Agent and Broker Liability: Crossing the Two Way Street*, 29 Ohio N.U. L. Rev. 237, 238 (2002) ("In many states, the distinction between 'agent' and 'broker' is at best a cloudy one.").

60. *See* Appleman, *supra* note 14, § 47.6, at 338.

61. *See, e.g.*, Douglas R. Richmond, *Insurance Agent and Broker Liability*, 40 Tort Trial & Ins. Prac. L.J. 1, 5 (2004) ("Brokers are sometimes referred to as independent agents and are generally considered to be the insured's agent." (citations omitted)). Richmond later qualifies this view, observing that "brokers" (as he defines them) may enter into "agency relationships" with insurance carriers and thus become "dual agent[s] for both the insured and the insurer. Specifically, a broker may be an agent of the insured for purposes of obtaining coverage and an agent of the insurer for other purposes." *Id.* at 7 (footnote omitted). Richmond's struggle will be familiar to anyone who has tried to survey the precise contours of the broker-agent distinction.

questions remain as to the precise legal status of such producers under agency law,[62] and to the scope of their legal duties to customers.

Legal distinctions aside, from a practical standpoint, it is generally recognized that brokers are most often used by large commercial customers, while smaller businesses and individuals purchase the bulk of their insurance through agents, either captive or independent.[63] Thus, in economic—if not in legal—terms, the insurance market is fairly clearly stratified. In fact, the two segments of the market have substantially different characteristics, the former tending toward oligopoly since the large broker consolidations of the 1990s[64] and the latter more resembling the "perfect competition" of economics textbooks with numerous firms contesting for relatively tiny slivers of the market.[65] We should not be surprised to discover that compensation mechanisms that are benign in the competitive agency market might be abused in the less competitive brokerage sector.

Some state regulators have been slow in recognizing the broker-agent distinction in the wake of the Spitzer Investigation, whose focus on the largest intermediaries, which conduct the bulk of their business as brokers (and in major historical "broker markets" like New York and San Francisco), obscured that issue for a time. As regulators outside New York begin to investigate second-tier insurance intermediaries, however, their efforts are being complicated by the fact that these intermediaries carry on a substantial portion of their business as independent agents, contractually bound to pursue the interests of their appointing carriers.[66] Indeed, in the first Spitzer Investigation-related complaint against an intermediary that explicitly addressed the broker-agent distinction, Connecticut's Attorney General, Richard Blumenthal, abandoned the breach of fiduciary duty thrust of Attorney General Spitzer's complaints (and his own) against Marsh, Aon, and Willis, presumably because the firm in question, Hilb Rogal & Hobbs, conducts the overwhelming majority of its business as an independent agent.[67] Instead, Attorney General Blumenthal alleged that the firm had violated Connecticut's unfair trade practices statutes by claiming to represent the interests of insureds despite its agency relationships with insurers.[68]

---

62. *See, e.g.*, Mate v. Wolverine Mut. Ins. Co., 592 N.W.2d 379, 386 (Mich. Ct. App. 1998) ("'Because brokers receive compensation from the insurers, it seems evident that a persuasive argument can be made for not treating a broker as an agent of the insurance purchaser.'" (quoting Robert E. Keeton & Alan I. Widiss, Insurance Law, A Guide to Fundamental Principles, Legal Doctrines, and Commercial Practices §2.5b(3), at 83-84 (1988))).

63. Cummins & Doherty, *supra* note 11, at 6-7.

64. *Id.* at 9.

65. *Id.* at 7-13.

66. *See* Complaint ¶¶ 4-15, Connecticut v. Hilb Rogal & Hobbs Co. (Conn. Super. Ct. Aug. 31, 2005), *available at* http://www.ct.gov/ag/lib/ag/consumers/hrh_complaint.pdf [hereinafter Hilb Rogal Complaint].

67. *Id.* ¶ 16.

68. *Id.* ¶¶ 16-22, 102-09.

So, while there may be some question as to the precise legal borders of the broker-agent distinction, it remains an important factor to be considered in assessing the appropriate response to the revelations of the Spitzer Investigation and its progeny.

B. *The Uses and Abuses of Contingent Commissions*

Insurance intermediaries are compensated in various ways for their work. In most cases, producers are compensated by the insurance carrier on a commission basis, although large corporate insurance buyers sometimes choose to compensate their brokers directly on a fee basis.[69] Commission-based compensation comes in two primary forms. The first is "standard commission" which is calculated as a percentage of the policy premium (typically ranging from five percent to twenty percent) and deducted by the producer from the customer's premium payment before it is remitted to the carrier. The second is "contingent commission," which is calculated as a percentage of a producer's entire book of business with a particular carrier (normally ranging from one percent to two percent) and paid by the carrier at year end based on the producer's meeting overall production and profitability goals.[70] Contingent commissions have been used by insurers as an incentive mechanism for their agents for a century or more.[71] Although contingent commissions have come under attack in recent months, their persistence can be attributed to substantial advantages they provide to insurers, intermediaries, and even consumers. But if, as I will argue, contingent compensation is the most efficient (e.g., low cost) means of promoting effective distribution of insurance products,[72] what risks does it pose and how can they be addressed?

To begin at the beginning, the affinity of insurers for contingent commissions as a mechanism for compensating producers is easy to grasp. It should resonate with anyone familiar with distributing products through

69. *See* Cummins & Doherty, *supra* note 11, at 14-16.

70. *Id.* at 14-15.

71. *See, e.g.*, Miss. Home Ins. Co. v. Adams & Boyle, 106 S.W. 209 (1907) (suit by an agent against a carrier to recover profit-based contingent commissions); Thompson v. Frelinghuysen, 191 Ill. App. 204 (1915) (suit by an agent of a brokerage firm to recover profit-based contingent commissions); W. Grain Dealers' Mut. Fire Ins. Co. v. Garrison Ins. Agency, 33 P.2d 950 (Kan. 1934) (holding that an agreement between an insurance agent and carrier provided for profit-based contingent commissions); Birdsall-Friedman Co. v. Globe & Rutgers Ins. Co., 190 A. 924 (Pa. 1937) (involving a profit-based contingent commission agreement). As early as 1951, the California Insurance Commissioner promulgated instructions to insurers with respect to their reporting in statutory filings of amounts of contingent commissions paid to agents. *See* Cal. Ins. Bulletin No. 113 (May 13, 1951).

72. Given that no fair observer can question that the P&C insurance market is, on the whole, a highly competitive one, it is fair to assume that compensation structures in the industry reflect perceived overall economies and a rational division of costs and benefits among the parties to insurance transactions. *See, e.g.*, Laureen Regan & Sharon Tennyson, *Agent Discretion and the Choice of Insurance Marketing System*, 39 J. Law & Econ. 637, 638 (1996) (noting that "different insurance marketing organizations arise as a means to minimize the costs of correctly matching policyholder risks with insurance coverage").

intermediaries, whether Oreos sold at Wal-Mart or Thin Mints sold by your friendly neighborhood Girl Scout. Simply put, distributors large and small will sell more of a particular product if they have incentives to do so, be those incentives volume bonuses or merit badges. In the case of commodities (i.e., fungible goods or services), the relative level of sales incentives may entirely dictate the focus of the distributor's energies between one similar product and another.

Insurance is, despite the best efforts of carriers to distinguish their products in the marketplace, substantially commoditized. Standard commissions, while variable and therefore susceptible to use as a sales incentive, are a blunt instrument and have no motivating effect beyond a single transaction. Moreover, standard commissions offered by one carrier in the context of a particular transaction will commonly be matched by one or more competitors eager to write the same account. Even if all competitors do not agree to pay higher commissions on any one transaction, the net effect of this type of competition over time is to increase commissions and, ultimately, the premiums paid by consumers above an optimum level. (And, of course, differential standard commissions have the same potential to induce producers to "steer" business to higher-paying insurers as contingents—a simple point, but one that has been largely overlooked in the debate over contingent compensation.) Contingent commissions, on the other hand, allow each carrier to compensate its producers on the basis of overall sales performance, fine-tuning the incentives it provides to its business strategy. While such compensation can also be the subject of "bidding up" by competing carriers, an independent agent or broker needs to maintain multiple sources of insurance capacity so it cannot leverage the "highest bidding" insurer on a book-of-business basis as it can in the context of a single transaction.

As importantly, typical contingent commission arrangements reward the intermediary not just for the volume of business produced, but for the profitability of that business. Although profit-based commissions have been used in the industry for many years, carriers have become steadily more committed to including this factor in their contingent compensation arrangements. Profitability is an important element in an effective incentive program because it puts the producer's "skin in the game," enlisting the agent or broker in the effort to identify customers that are less likely to incur insured losses. Such an alignment of interests has benefits for consumers, as well as providers, of insurance, as I discuss below. At a minimum, such profit sharing provides a disincentive for the producer to attain its sales goals by loading up the carrier with substandard risks.[73]

73. Given insurers' strong interest in linking producer compensation to profitability, Marsh's success in the late 1990s in eliminating profitability as an element in its contingent compensation through the device of the PSA speaks volumes about the relative bargaining power of that mega-broker and the many competing insurers for which it places business. J. David Cummins and Neil A. Doherty, in their recent study of the economics of insurance

Interestingly, given the financial stakes involved, there has been little independent research indicating that contingent commissions are effective in influencing producer behavior. The mere fact that this device has been so widely adopted by insurers over the past century is probably proof enough of its efficacy, however.[74] The one academic study of the topic, by an MIT doctoral candidate named Jeffrey Wilder, found that contingent commissions do indeed influence the behavior of individual producers who are aware of them (i.e., equity owners in an insurance agency).[75] Wilder also noted that contingent commissions can theoretically influence the behavior of nonowner producers as well, if an agency creates internal incentives for them in support of the firm's goals.[76]

If the appeal of contingent commissions to insurance carriers is fairly clear, why has this form of compensation been embraced by the producers themselves? Why would they not simply insist on higher standard commissions on the front end and eliminate the risk that bad results on their placements would reduce their compensation? The simplest answer to this question is that insurance agents and brokers need access to insurance carriers as much as the carriers need access to customers, and the market for insurance placement services remains more than competitive enough to prevent individual intermediaries from exacting standard commissions equivalent to what carriers are willing to pay on a contingent basis to their most successful producers.[77]

There is perhaps another, more subtle, reason that contingent commissions have developed as a compensation technique distinct from standard commissions, a reason having to do with the internal economics of insurance agencies. Most individual insurance agents or brokers are employees of their firms, with little or no equity stake in the business. They do, however, "own" their books of business to a large extent, giving successful agents considerable bargaining power vis-à-vis their employers. Insurance brokerage is an "eat what you kill" business, meaning that the bulk of an individual agent's compensation is based on a percentage of the

---

intermediaries, note the potential for abuse in volume-only contingent commission arrangements. Cummins & Doherty, *supra* note 11, at 16-18.

74. At a more fundamental level, Laureen Regan and Sharon Tennyson have demonstrated that insurers will gravitate to independent agent and broker distribution, as opposed to direct sales or captive agent distribution, for more complex products where the intermediary's superior access to information concerning the risk will produce improvements in underwriting results outweighing any increased compensation costs involved. *See* Regan & Tennyson, *supra* note 72, at 646-47.

75. Jeffrey Wilder, *Competing for the Effort of a Common Agent: Contingency Fees in Commercial Insurance* 21 (U.S. Dep't of Justice Antitrust Div. Econ. Analysis Group, Working Paper No. EAG03-4, 2004), *available at* http://papers.ssrn.com/sol3/papers.cfm?abstract_id=418061. An earlier version of Wilder's paper was cited in the WLF's February 10, 2004, letter that was the initial catalyst of the Spitzer Investigation. WLF Letter, *supra* note 7, at 2.

76. Wilder, *supra* note 75, at 21.

77. *See generally* Cummins & Doherty, *supra* note 11, at 10-13.

standard commission income she brings into the firm.[78] Accordingly, an agency principal would have a strong incentive to reduce the firm's standard commission income if such income could be replaced with revenue that was less visible to employee agents—like year-end contingent commission payments calculated as a percentage of the firm's entire book of business. Indeed, it may be that efforts to preserve the confidentiality of contingent commission arrangements have had less to do with keeping such information from an insurance agency's customers than they have with keeping such information from the agency's own employees. While agency owners are understandably reluctant to concede this for the record, many will state in private that contingent commissions provide a vital pool of "house" money that funds the basic overhead of their firms. Off the record, many principals of smaller agencies question whether they could remain in business without contingent compensation reserved to the firm.[79] Such compensation also provides an incentive for an insurance agency to guard its firm-wide reputation for ethical conduct (without which it is likely to lose agency appointments from carriers), and should therefore improve oversight of individual producer's sales practices.

The potentially catacylsmic impact on the level of competition in the market for insurance intermediation services of an outright ban on contingent commissions has gone largely unnoted in the debate spawned by the Spitzer Investigation. It may well be, however, that Attorney General Spitzer's concession at the National Press Club that contingent commissions can be appropriate in some circumstances indicates that he understands this risk.

But what of the customers? Are the interests of insurance consumers prejudiced by contingent compensation agreements between insurers and intermediaries? Attorney General Spitzer and others have alleged that such compensation artificially raises the price of insurance, as opposed to reflecting its true costs. It is not at all clear, however, that this is the case. To the contrary, a strong case can be made that independent intermediaries compensated in part based on the performance of their overall books of business with individual insurance carriers are the most efficient distribution mechanism for complex insurance products.[80] Further, contingent commissions may generate benefits to insureds—in terms of the "exposure cost" (that is, the risk-based premium net of expenses) and availability of coverage—in addition to reducing the transaction costs of the insurance purchase.

From the standpoint of basic economics, the mere durability of the independent broker or agent distribution model as a feature of the market for certain kinds of insurance products argues for its efficiency.

78. *See* Wilder, *supra* note 75, at 7-8.

79. *See, e.g.*, Phil Zinkewicz, *Agents Being Swept into the Scandal*, Rough Notes, Dec. 2004, at 84, 84-86.

80. *See, e.g.*, Cummins & Doherty, *supra* note 11, at 17-18; Regan & Tennyson, *supra* note 72, at 645-46.

Notwithstanding market distortions such as Attorney General Spitzer apparently discovered in the Excess Casualty area at Marsh, the broader P&C insurance market is unquestionably a truly competitive one, with low barriers to entry and an array of available distribution channels all but guaranteeing that consumers will in most circumstances have access to a desired mix of insurance products and services at costs free of what economists refer to as "monopoly rents."[81] Indeed, Laureen Regan and Sharon Tennyson have demonstrated that insurers will gravitate to lower cost distribution structures such as direct sales or captive agents in lines where underwriting is relatively generic and claims costs more predictable (e.g., personal lines coverages such as auto), that is, where they can compete for profitable business without incurring, and passing on to the consumer in whole or in part, the higher costs of an independent intermediary.[82] And, if the P&C insurance market is fundamentally competitive, then consumers must benefit from the efforts of providers and intermediaries to achieve competitive advantage in terms of cost and other factors. At a minimum, therefore, the current independent producer marketing organization is due a presumption of economic legitimacy.

With specific regard to profit-based contingent commissions, some consumers (specifically, good risks) will also benefit from lower prices because such commissions provide an incentive for producers to provide sufficient information to the carrier to overcome its concern with the problem of "adverse selection," which arises because an insured inevitably knows more about its true risks than the insurance carrier. As J. David Cummins and Neil A. Doherty have observed,

> Insofar as the insurer is unable to differentiate risks, it may try to charge uniform premiums to good (low) risks and bad (high) risks alike. Thus, the good risks will end up subsidizing the bad. This implies that bad risk

81. *See* Cummins & Doherty, *supra* note 11, at 7-14. It may be argued that this view of the insurance market is overly simplistic, and in particular that the market for insurance intermediation services demonstrably has sectors that are less than perfectly competitive. This argument is probably strongest with respect to two sectors at the extremes of the commercial P&C market: commercial insurance for large corporations on the one hand and personal lines insurance for middle-income families on the other. The former sector sees limits on competition for brokerage business because the enormous resource requirements of large commercial can only be met by a small number of large firms. At the smaller end of the spectrum, competition may be limited in practical terms by the reluctance of consumers to "shop" their business from one agent to another for reasons of loyalty or simple aversion to change. While such arguments cannot be dismissed out of hand, I would argue that even in these sectors, aggressive competition does occur—whether in the form of Aon proposing alternative programs to Marsh clients in the Fortune 500 or a direct marketer of personal auto insurance sending an unsolicited rate quote to an individual consumer. Thus, even if buyers of insurance in some quarters, large and small, show reluctance to change intermediaries frequently, they nonetheless have access to market information enabling them to assess whether they are getting a "good deal" through their chosen intermediary.

82. *See* Regan & Tennyson, *supra* note 72, at 645-46. Cummins's and Doherty's research indicates that producer compensation is passed through to consumers in increased prices, but not fully, so insurers retain an economic interest in achieving the lowest priced distribution model that will produce a sustainable book of business, as of course, do consumers. *See* Cummins & Doherty, *supra* note 11, at 19-21.

> policyholders will find insurance very attractive and will demand considerable insurance, but the demand for insurance by good risks will be light and might disappear altogether. Thus, asymmetric information "crowds out" the good risks, and insurance is only fairly priced for the bad risks. The insurance market ends up with an adverse selection problem, with the insured population representing primarily the higher risk clients. The costs of adverse selection fall on policyholders, particularly the good risks, who may have to pay excessive rates or accept diminished coverage. Insurers recognize the problems caused by lack of information, and this is reflected in the prices and coverage they offer. But, if adverse selection can be avoided, policyholders will be better off.[83]

Thus, contingent compensation arrangements that provide incentives to producers to assist in overcoming the adverse selection problem will benefit policyholders by encouraging insurers to price good risks more competitively. Transaction costs can also be lowered through this device, because the costs of an independent agent's or broker's vetting of a risk are in the aggregate spread among a number of carriers to which it might submit business. Finally, profit-based contingents give producers a significant incentive to provide their clients with loss control services after they have sold a policy—and help fund such services.

Contingent commissions should arguably be preferred over standard commissions by consumers for another reason: Higher up-front commissions encourage a short-term focus by producers, reducing their reluctance to place business with markets irrespective of their long-term prospects. Such a short-term focus, and the periodic market contractions it contributes to as marginal carriers fail, exacerbates the volatility of the underwriting cycle and ultimately hurts consumers.[84] Contingent commissions, on the other hand, give the producer a stake in making sure that a carrier remains solvent and able to pay claims, as well as annual performance awards to its producers.

Implicit in Attorney General Spitzer's critique of contingent commissions, however, is a belief that—whatever market-proven efficiencies they may represent and whatever benefits to consumers they may provide—such compensation is a corrupting influence on producers because of the temptations it provides to "steer" business in order to maximize benefits to the producer as opposed to the customer. This concern is not utterly unfounded, as no one would deny that producer incentives are designed by carriers to attract business, but it overlooks one fundamental fact. Insurance intermediaries are normally more concerned with the risk of losing a good customer to a competing producer than they

83. Cummins & Doherty, *supra* note 11, at 25. Even contingent commissions based on volume alone may have benefits for consumers, ranging from lowered costs achieved by economies of scale within insurers to increased competition from new markets using such commissions to gain market share. *Id.* at 17, 24.

84. Fitzpatrick, *supra* note 14, at 269-70.

are with any marginal inducements that may be provided by any one insurance carrier.[85] A typical agent or broker is therefore extremely unlikely to act in such a way that it will risk losing a customer and the long-term income stream it represents, whatever immediate benefit it might derive in contingent compensation.[86] Even less compelling is the claim made by some critics that profit-based contingent commissions will cause producers to be less than aggressive in seeking payment for their client's claims.[87] 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; the risk to individual customer relationships and to the intermediary's reputation in the marketplace would both militate strongly against a contingent commission-induced effort to reduce claims collections.[88]

If concerns with endemic market manipulations driven by contingent commission arrangements are largely unfounded, however, a belief that the insurance market is fundamentally competitive has as its necessary corollary the acknowledgement that greater transparency in producer compensation would not alter the market's essential structure and operations. This, in truth, is the challenge that has been thrown down by Attorney General Spitzer: If the insurance market is efficient and serves consumers well, what have its participants to fear from a little sunshine?

But before we explore the implications of this challenge, let us first examine the early legislative and regulatory responses to Spitzer's challenge.

III. LEGISLATIVE AND REGULATORY RESPONSES: THE SMALL LAWS

By mid-2005, the focus of regulatory debate on the implications of the Spitzer Investigation had shifted from state attorneys general ("AGs") to state legislators and insurance commissioners. This shift was not attributable to any lack of continuing prosecutorial zeal on the part of state AGs, as demonstrated by ongoing civil and criminal investigations on multiple fronts, but instead reflected the real limitations of their power to drive prospective market reform other than by consent agreements with individual defendants. One can also speculate that state prosecutors have been reluctant to attack the well-established institution of contingent

85. As noted above, I have explored the motivations of insurance producers and other insurance market participants in more detail in an earlier work. *See id.* at 269 ("Whatever an agent's or broker's legal relation to the underwriting carrier they are motivated, in practical terms, by the fear that they will lose their customer to another agent or broker who can deliver the same coverage at a lower price."). Of course, circumstances can exist—as they appear to have existed in the Excess Casualty unit at Marsh Global Broking—where individual producers can be simultaneously insulated from the salutary effects of this fear and motivated by other incentives. Such combinations of circumstances remain rare, however.

86. *See, e.g.*, Cummins & Doherty, *supra* note 11, at 23; Wilder, *supra* note 75, at 7.

87. *See, e.g.*, Hunter, *supra* note 33.

88. *See* Cummins & Doherty, *supra* note 11, at 27-28.

compensation for independent agents, as opposed to brokers, either because they appreciate the risk of unintended harm to an important "Main Street" business sector or because they fear alienating a powerful political constituency.[89]

Legislative and administrative market reform efforts have targeted disclosure of compensation received by insurance producers, as opposed to proscribing contingent commissions themselves. Two organizations of state insurance officials, the National Association of Insurance Commissioners ("NAIC") and the National Counsel of Insurance Legislators ("NCOIL"), have promulgated model producer disclosure legislation in response to the Spitzer Investigation. The NAIC acted first, in December 2004. The initial NAIC draft addressed producer disclosure at two levels. Section A of the draft prohibited any producer that "receives any compensation from the customer . . . or represents the customer" in an insurance placement from accepting compensation from the carrier unless the customer has consented in writing and been informed of the amount of such compensation.[90] Section B of the original NAIC draft also contained a "generic disclosure" provision, which provided that—irrespective of whether a producer was being paid by or acting on behalf of the customer—the producer must disclose to its customer that it will receive compensation from the carrier writing the customer's policy, that such compensation might vary from carrier to carrier, and that the producer might receive compensation from the carrier based on aggregate performance of its book of business with that carrier.[91] Ultimately, however, the NAIC adopted a

---

89. While not as numerous as lawyers, insurance agents are a substantial bloc in state legislatures. *Cf.* Peverill Squire, *Legislative Professionalization and Membership Diversity in State Legislatures*, 17 Legis. Stud. Q. 69, 74-75 (1992).

90. In relevant part, Section A of the NAIC Model Act provides as follows:

> A.(1) Where any insurance producer or any affiliate of such producer receives any compensation from the customer for the placement of insurance or represents the customer with respect to that placement, neither that producer nor the affiliate shall accept or receive any compensation from an insurer or other third party for that placement of insurance unless the customer has, prior to the customer's purchase of insurance:
>
> (a) Obtained the customer's documented acknowledgement that such compensation will be received by the producer or affiliate; and
>
> (b) Disclosed the amount of compensation from the insurer or other third party for that placement. If the amount of compensation is not known at the time of disclosure, the producer shall disclose the specific method for calculating such compensation and, if possible, a reasonable estimate of such amount.

Proposed Compensation Disclosure Amendment to the Producer Licensing Model Act (Nat'l Assoc. of Ins. Comm'rs 2004) [hereinafter NAIC Model Act], *available at* http://www.naic.org/documents/committees_ex_broker_comp_disclosure.pdf (version adopted Dec. 29, 2004).

91. The deferred Section B provided, in pertinent part as follows:

> B. An insurance producer must disclose the following, if applicable, to a customer, prior to the purchase of insurance:
>
> (1) That the producer will receive compensation from an insurer or other third party for the sale;
>
> (2) That the compensation received by the producer may differ depending upon the product and insurer(s); and

model disclosure provision that contained only Section A, and deferred consideration of the broader requirement of draft Section B.

NCOIL took an even more conservative approach, adopting a model disclosure law in March 2005 that requires a producer to obtain consent from and disclose its compensation to the customer only where the producer will be paid by both parties to the transaction.[92] In other words, the NCOIL model did not include the alternative basis for requiring disclosure contained in Section A of the NAIC model: that the producer "represents the customer with respect to the placement." The NCOIL model also limited its requirements to a customer's "initial" placement of insurance (as opposed to renewals) and reduced the specificity required in disclosing formula-based contingent compensation.[93]

In the months since the NAIC and NCOIL models were issued, states that have acted in this area have shown an almost uniform preference for the NCOIL model. At this writing, a total of seven states have enacted new producer disclosure laws; none has adopted the NAIC model outright and only two have adopted requirements beyond those contained in the NCOIL model.[94] Significantly, neither New York nor California has yet addressed broker disclosure legislatively. New York seems unlikely to enact broad insurance market reform legislatively. Indeed, notwithstanding New York Attorney General Spitzer's attacks on contingent commissions in general,

---

> (3) That the producer may receive additional compensation from an insurer or other third party based upon other factors, such as premium volume placed with a particular insurer and loss or claims experience.

NAIC Model Act, *supra* note 90, *available at* http://www.naic.org/documents/committees_ex_broker_Broker_subsecB.pdf (draft of Dec. 27, 2004).

92. The NCOIL Model Act provides, in pertinent part as follows:

> A. Where any insurance producer or any affiliate of such producer receives any compensation from the customer for the initial placement of insurance, neither that producer nor the affiliate shall accept or receive any compensation from an insurer or other third party for that placement of insurance unless the producer has, prior to the customer's purchase of insurance:
>
> (1) Obtained the customer's documented acknowledgement that such compensation will be received by the producer or affiliate; and
>
> (2) Provided 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 (Nat'l Conference of Ins. Legislators 2005) [hereinafter NCOIL Model], *available at* http://www.aba.com/aba/documents/abia/NCOILFinalMarkup.pdf.

93. *Id.*

94. Connecticut, Georgia, Oregon, and Texas have enacted versions of the NCOIL Model. Arkansas and Rhode Island have enacted similar legislation with an additional requirement that all producers receiving compensation from the insurer disclose that fact. Nevada's Division of Insurance has issued a temporary rule imposing substantial disclosure requirements on brokers, but exempting agents, defined as those producers that are "compensated by the insurer." State of Nev., Dep't of Bus. and Indus. Div. of Ins., Temporary Regulation Concerning Broker's Duties to Client: Duty Against Self-Dealing, Duty to Disclose Compensation and Duty to Disclose All Quotes; Violations (Feb. 3, 2005), *available at* http://doi.state.nv.us/Laws-REG-Temp-Broker-05.pdf.

the state's insurance superintendent has said publicly that independent agents, as opposed to large brokers, should be entitled to accept contingent commissions.[95] California's insurance commissioner has taken a more aggressive stance, proposing broad-based regulations imposing new duties on insurance brokers and agents.[96] There is a substantial question, however, whether Commissioner John Garamendi has the statutory power to impose such requirements under existing California law,[97] and the California legislature has thus far rebuffed his attempts to obtain such authority through new legislation.[98]

95. *See* Michael Ha, *NY's Mills Defends Fees for Independent Agents*, Nat'l Underwriter Apr. 29, 2005, http://www.nationalunderwriter.com/pandc/hotnews/viewPC.asp?article=4_29_05_14_17204.xml&src=5.

96. *See* Press Release, Cal. Dep't of Ins., Insurance Commissioner John Garamendi Unveils New Regulations in Ongoing Battle to Protect Consumers from Impact of Secret Broker Commissions (Apr. 13, 2005), *available at* http://www.insurance.ca.gov/0400-news/0100-press-releases/0080-2005/release039-05.cfm. Commissioner Garamendi's proposals are by far the most far reaching seen to date. Specifically, Commissioner Garamendi's proposed regulations would require the following:

> Require the agent or broker to advise a client, prior to signing an agreement or receiving a fee, whether the producer will seek a quote from one insurer or more than one insurer.
>
> Require the agent or broker to reveal if he or she is acting on behalf of the insurer or the client in connection with the placement of insurance. A producer who accepts a fee from a client is conclusively deemed to be acting on behalf of the client.
>
> Require the agent or broker to reveal the amount of compensation he or she will receive if the client purchases insurance with any insurer recommended by the agent or broker. If the amount of compensation cannot reasonably be known at the time this disclosure is made, the producer may disclose the method by which any such compensation will or may be calculated.
>
> A broker or agent acting on behalf of the client, or who accepts a fee from the client, may not accept any compensation from a third party for the transaction done on behalf of the client without first obtaining the consent of the client.
>
> A broker or agent who has told a client that he or she will search for the best quote on a policy must reveal the number of quotes obtained, the name of the insurer, the premium amount, and other required information.

Id. Commissioner Garamendi's April 2005 proposed regulations were less far reaching, however, than provisions he proposed in October 2004 in the immediate aftermath of New York Attorney General Spitzer's suit against Marsh. *See, e.g.*, Lord, Bissell & Brook, Client Alert, California Proposes Revamped Insurance Producer Disclosure Regulations (Apr. 22, 2005), http://www.lordbissell.com/Newsstand/2005-04_RevisedCABrokerRegs_Barney.pdf. As of this writing, Commissioner Garamendi has postponed issuing the regulation he proposed in April 2005, citing a move toward voluntary disclosure of agency and compensation arrangements being promoted by a leading producer trade group. *See infra* note 110.

97. *See State Fund's Policy May Set Precedent, Perhaps a Bad One*, Workers' Comp Executive, Sept. 15, 2005, http://www.wcexec.com/vdata/5/nl/12/everything/wce_esv_05i14_015_article2.htm.

98. *See* News Release, Nat'l Ass'n of Mutual Ins. Cos., California Legislative Committee Takes Correct Action on Producer Requirements (May 2, 2005), http://www.namic.org/newsreleases05/050502nr1.asp. On September 30, 2005, the California Department of Insurance again entered the mix, issuing an informal letter opinion to "explain its legal position on the fiduciary duties of brokers and independent agents under California law." Letter from Jon A. Tomashoff, Senior Staff Counsel, Cal. Dep't of Ins., to Stephen Young, Esq., IBA West (Sept. 30, 2005) (on file with author). Not surprisingly, the

Thus, a full year after Attorney General Spitzer shocked the insurance world with his bid-rigging allegations against Marsh, the "business reforms" settlements he has imposed on Marsh, Aon, and Willis remain the most substantial new market regulation—in practical terms—yet effected. Legislative responses thus far have been few, and those that have been enacted are quite limited in scope. These "small laws" tell us something about both the elected officials who enacted them and the public, which thus far seems satisfied with them: that both understand (intuitively, at least) that the market dysfunctions which gave rise to abuses at large global insurance brokers like Marsh, Aon, and Willis have little relevance to the independent insurance agent on Main Street, U.S.A., or to his customers.

But where does this leave us? How are we to resolve the very real issues raised by the Spitzer Investigation about the opacity of the insurance market to the consumers it serves? How are policyholders to be assured that their intermediaries are properly managing the potential conflicts of interest inherent in their function? Put in more prosaic terms, how and on what basis is the Spitzer Investigation to be resolved? What terms of armistice will be acceptable to the stakeholders in this controversy, be they prosecutors, legislators, insurance intermediaries, carriers, or consumers? In the next section, I outline one way to achieve such a concord: a simple new "social contract"[99] for the P&C insurance market.

IV. A MODEST PROPOSAL: INSURANCE "IN THE SUNSHINE"

The foregoing has, I hope, demonstrated two things. First, the simplistic invocation of loaded terms like "kickback" fails to capture the complexity of the economic and legal relationships that have developed over centuries in the insurance market, and one would disturb that edifice of relationships at some risk to the interests of insurance consumers. Second, notwithstanding the above observation, Attorney General Spitzer is on to something: He has detected a reluctance in insurance practitioners to disclose the details of producer compensation that would seem to belie the industry's assertion that its prevailing distribution represents the best of all

Department's counsel asserts that both brokers and independent agents (when acting as "dual agents" for both the carrier and the insured) owe fiduciary duties to customers, despite California case law indicating that the producer-insured relationship is something less than a fiduciary one. *See, e.g.*, Hydro-Mill Co. v. Hayward, Tilton & Rolapp Ins. Assocs., 10 Cal. Rptr. 3d 582 (Ct. App. 2004); Kotlar v. Hartford Fire Ins. Co., 100 Cal. Rptr. 2d 246 (Ct. App. 2000).

99. Thomas Donaldson & Thomas W. Dunfee, *Précis for Ties that Bind*, 105 Bus. & Soc'y Rev. 436, 442 (2000). Donaldson and Dunfee believe that

> the same logic that sanctifies a handshake between two individuals turns out also to sanctify the implicit understandings of economic communities woven throughout the business world. These are the informal but critical agreements—or "social contracts"—that provide the warp and woof of economic life. These are the agreements that exist within industries, national economies, trade groups, and corporations, and, further, that are the implicit "contracts" critical for understanding business ethics.

Id.

possible worlds for consumers. Given this reluctance, one could forgive the Attorney General for thinking that the insurance industry "doth protest too much" in reciting the economic case for contingent commissions. And, while Spitzer has, through the medium of private settlements, largely cured the market dysfunctions born of the excessive concentration of the brokerage market for large corporate risks that occurred in the 1990s, he plainly has lingering concerns about the rest of the P&C market. What, then, would ease these concerns and bring an end to the Spitzer Investigation and its progeny in other states?

Transparency, my friends, transparency: This is the *sine qua non* of a peace treaty between Attorney General Spitzer and the insurance industry. There are models aplenty of how this could be accomplished. Some U.S. insurers have, for example, voluntarily placed information regarding their producer compensation practices on their public web sites.[100] This approach has been endorsed by the largest organization of independent agents, the Independent Insurance Agents & Brokers of America ("IIABA"), better known as the Big "I."[101] But as helpful as disclosure by insurers can be, most consumers' primary point of contact with the insurance market will continue to be agents and brokers. Any legal duties of disclosure that exist also reside in that relationship, as opposed to the more remote relationship between insured and insurer. Should insurance producers, in addition to carriers, be encouraged to disclose the basis of their compensation?

Real estate agents and brokers, for example, have in recent years adapted to a system in which a consumer cannot even be shown a property without first deciding how the intermediary will be compensated and signing an acknowledgement that the consumer understands who will pay the intermediary and whose interests the intermediary represents. The real estate industry is a particularly useful analog for purposes of considering reform in the insurance market. An intermediary-driven market like insurance, the real estate market has been regulated similarly over the past century—with every U.S. state maintaining a "licensing statute or regulatory scheme addressing qualifications for obtaining the necessary real estate salesperson's or broker's license, and regulations governing realtors'

100. The most comprehensive such disclosure is The Chubb Corporation's, *available at* http://www.chubb.com/marketing/chubb3887.html (last visited Apr. 17, 2006); the author participated in the development of this disclosure. Other major carriers such as The Hartford, *available at* http://www.thehartford.com/servlet/Satellite?cid=1122655319479&pagename=HIG/Page/PopUp&nt_page_id=1122655319479&nt_section=1122655319479&c=Page (last visited Apr. 17, 2006), and St. Paul Travelers, *available at* http://www.stpaultravelers.com/legal/producerCompDscl.html (last visited Apr. 17, 2006), have also posted producer compensation disclosures.

101. Press Release, Indep. Agents & Brokers of Am., Big "I" Board Adopts Policy on Insurance Company Disclosure (Sept. 23, 2005), *available at* http://www.iiaa.org/eprise/main/CB_Website/Affiliated/NationalAssociation/IIAA/02_News/02_PressRelease/NA20050923143814.

activities and conduct."[102] Like today's insurance regulations, however, traditional state realtor licensing statutes did not typically mandate particular agency, compensation, or disclosure arrangements, which instead developed primarily in response to the particular economics of the real estate purchasing process.[103]

The real estate industry's wake-up call came in 1983, when a Federal Trade Commission study found that seventy-two percent of all home buyers believed, erroneously, that the "selling agent" assisting them was representing their interests.[104] In fact, under long-established industry practices, such intermediaries were sub-agents of the seller of the property. In the ensuing decades, most states responded by mandating disclosure to consumers of such agency relationships.[105] Many states themselves provide consumers with information on real estate agency relationships.[106] One feature of virtually all such disclosure regimes is the acknowledgment of "dual agency": That is, an intermediary may act for both sides to a real estate transaction so long as this dual agency is disclosed.[107]

Such disclosures as are now mandated in the real estate industry have the benefit of being simple and not providing so much information as to be useless to consumers in practical terms.[108] For example, one real estate

---

102. Ann Morales Olazábal, *Redefining Realtor Relationships and Responsibilities: The Failure of State Regulatory Responses*, 40 Harv. J. on Legis. 65, 70 (2003).

103. *Id.* at 71 ("State licensing statutes did not, however, dictate the form of agency representation then prevalent. Instead, the entrenchment of the listing/cooperating or 'traditional' agency representation model was a direct result of the multiple listing systems in use nationwide.").

104. L.A. Reg'l Off., Fed. Trade Comm'n, Residential Real Estate Brokerage Industry 1, 69 (1983). In what perhaps should stand as an object lesson to the insurance industry, the real estate industry did not immediately embrace changes to the traditional market models proposed in the wake of the 1983 Federal Trade Commission study. Indeed, it was not until 1992 that the National Association of Realtors ("NAR"), under political pressure from consumer groups and economic pressure from a new sector of "exclusive buyer's agents," agreed to eliminate the requirement that agents accessing its regional and local multiple listing services be subagents of sellers. Olazábal, *supra* note 102, at 74-75. This voluntary reform recognized, perhaps belatedly, that change in the real estate market was inevitable, and allowed the NAR to retain the initiative in guiding the evolution of real estate market practices.

105. *See* Craig W. Dallon, *Theories of Real Estate Broker Liability and the Effect of the 'As Is' Clause*, 54 Fla. L. Rev. 395, 416-18 (2002).

106. *See, e.g.*, N.Y. State Dep't of State, Div. of Licensing Servs., Disclosure Regarding Real Estate Agency Relationships (2001), *available at* http://www.dos.state.ny.us/lcns/pdfs/dos1565.pdf; State of Cal. Dep't of Real Estate, Disclosures in Real Property Transactions (1999), *available at* http://www.dre.ca.gov/disclosures.htm#agency.

107. Olazábal, *supra* note 102, at 79.

108. The "privacy policy" disclosures required of financial institutions by the Gramm-Leach-Bliley Act of 1999, 15 U.S.C. 6801-6809 (2000), come to mind as an example of disclosure "overkill." Does any appreciable portion of the consuming public actually read these impenetrable missives? There are, of course, commentators who believe that current real estate disclosures are insufficient, as there would no doubt be those who make the same critique of my proposed insurance producer disclosure. Interestingly, however, the complaint lodged against current realtor disclosures is that they are too complex to be understood by unsophisticated homebuyers, not that they are too simple. *See, e.g.*, Olazábal,

firm in my state, Connecticut, provides its customers with a simple "Consumer Information Statement of Real Estate Relationships" that outlines the various ways in which an intermediary can participate in a transaction:

1. AS A SELLER'S AGENT . . . I AS A LICENSEE, REPRESENT THE SELLER AND ALL MATERIAL INFORMATION SUPPLIED TO ME BY THE BUYER WILL BE TOLD TO THE SELLER.

2. AS A BUYER'S AGENT, I . . . REPRESENT THE BUYER AND ALL MATERIAL INFORMATION SUPPLIED TO ME BY THE SELLER WILL BE TOLD TO THE BUYER.

3. AS A DISCLOSED DUAL AGENT, I . . . REPRESENT BOTH PARTIES. HOWEVER, I MAY NOT, WITHOUT EXPRESS PERMISSION, DISCLOSE THAT THE SELLER WILL ACCEPT A PRICE LESS THAN THE LISTING PRICE OR THAT THE BUYER WILL PAY A PRICE GREATER THAN THE OFFERED PRICE.[109]

This disclosure is brief, straightforward, and gives the client the option of choosing how he or she wishes to work with the intermediary. It is simple enough to be understood by an unsophisticated purchaser of real estate, while being sufficient to initiate a more far-reaching dialogue with even the most sophisticated.

I suspect that Mr. Spitzer and his colleagues in state AGs' and insurance commissioners' offices around the country would consider themselves to have achieved substantial market reform, perhaps to the point of resolving their broader concerns over producer compensation, if the Big "I" and other independent producers' organizations would propose—and their members would voluntarily adopt—a disclosure form for new customers that sets out the basic facts of the intermediary's agency relationships. Such a disclosure might look like this:

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.

---

supra note 102, at 123-24. This criticism in itself would seem to support the idea that a simple, high-level approach to disclosure will benefit the greatest number of consumers.

109. Fairfield County Homes, Connecticut Consumer Statement on Real Estate Relationships (2005), *available at* http://www.fairfieldcountyhomes.info/ConsumerStatementonRealEstateRelationships.

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.

IF YOU CHOOSE, WE WILL REPRESENT YOUR INTERESTS EXCLUSIVELY AS A "BROKER" IN YOUR INSURANCE PURCHASE. IF WE ACT AS YOUR "BROKER," WE WILL BE COMPENSATED BY YOU, WITH OUR FEE TO BE DETERMINED ACCORDING TO THE ATTACHED SCHEDULE, AND WE WILL ACCEPT NO COMMISSION OR ADDITIONAL COMPENSATION FROM AN INSURER IN CONNECTION WITH YOUR PURCHASE.

Adoption of a disclosure form along these lines would probably change few consumers' approach to purchasing insurance; most would presumably prefer to trade on a commission basis, just as most home buyers continue to purchase real estate through agents compensated by the seller. Such an approach would add little extra time or inconvenience to the insurance purchasing process, and any new costs imposed would be de minimis. No complex calculations would be necessary; no estimates of contingent payments due, if at all, only after a review of an entire year's performance, would be required. Most importantly, this reform would preserve important mechanisms by which insurance carriers and their appointed agents have sought for a century or more to align their interests and control their costs, while respecting the consumer's right to understand the fundamental legal and economic relationships underlying his or her insurance purchase.[110]

Clearly, a voluntary solution along these lines would be preferable to the ongoing costs in time, resources, and distraction of the ongoing insurance market investigations. It would recognize the time-honored economic value of the independent agency system, preserve its efficiencies, and remove the cloud of suspicion raised by Attorney General Spitzer's discovery of real abuses involving large brokerage firms whose control of a few niches of the

110. On October 20, 2005, the Insurance Brokers & Agents of the West ("IBA West"), a leading producer trade group on the West Coast, issued to its members the IBA West Guide to Compensation Disclosure. While not taking a formal position on disclosure by its members of agency relationships and compensation structures, the Guide offers model disclosure language in a variety of forms and identifies "considerations for members to evaluate should they decide, in light of their own professional circumstances, to provide more information to their customers." *See* IBA West, IBA West Guide to Compensation Disclosure (2005), *available at* http://www.ibawest.com/pdf/Articles/IBAWestDisclosureGuide110105Final.pdf; Letter from Stanley Simpson, President, IBA West, to IBA West Members 2 (Oct. 20, 2005), *available at* http://www.ibawest.com/pdf/Articles/CompensationDisclosureCoverLetterGuide110105.pdf. Even this careful nod toward voluntary disclosure generated a ringing endorsement from California Insurance Commissioner John Garamendi, who announced that, in light of the issuance of the Guide he would hold off on issuing his proposed regulation imposing additional duties on California producers. *See supra* note 96; *see also* Press Release, Cal. Dep't of Ins., Insurance Commissioner John Garamendi Commends IBA West for Its Guide to Agent-Broker Compensation Disclosure (Nov. 2, 2005), *available at* http://www.insurance.ca.gov/0400-news/0100-press-releases/0080-2005/release-100-05.cfm.

insurance market grew to the point where normal checks (primarily fear of competition) on an otherwise benign system of producer compensation ceased to operate effectively. If adopted, the independent agency system, and the insurers and consumers who rely upon it, could return to the business of efficiently spreading society's risks for the benefit of all. Sunshine, as Louis Brandeis observed, is the best disinfectant.[111]

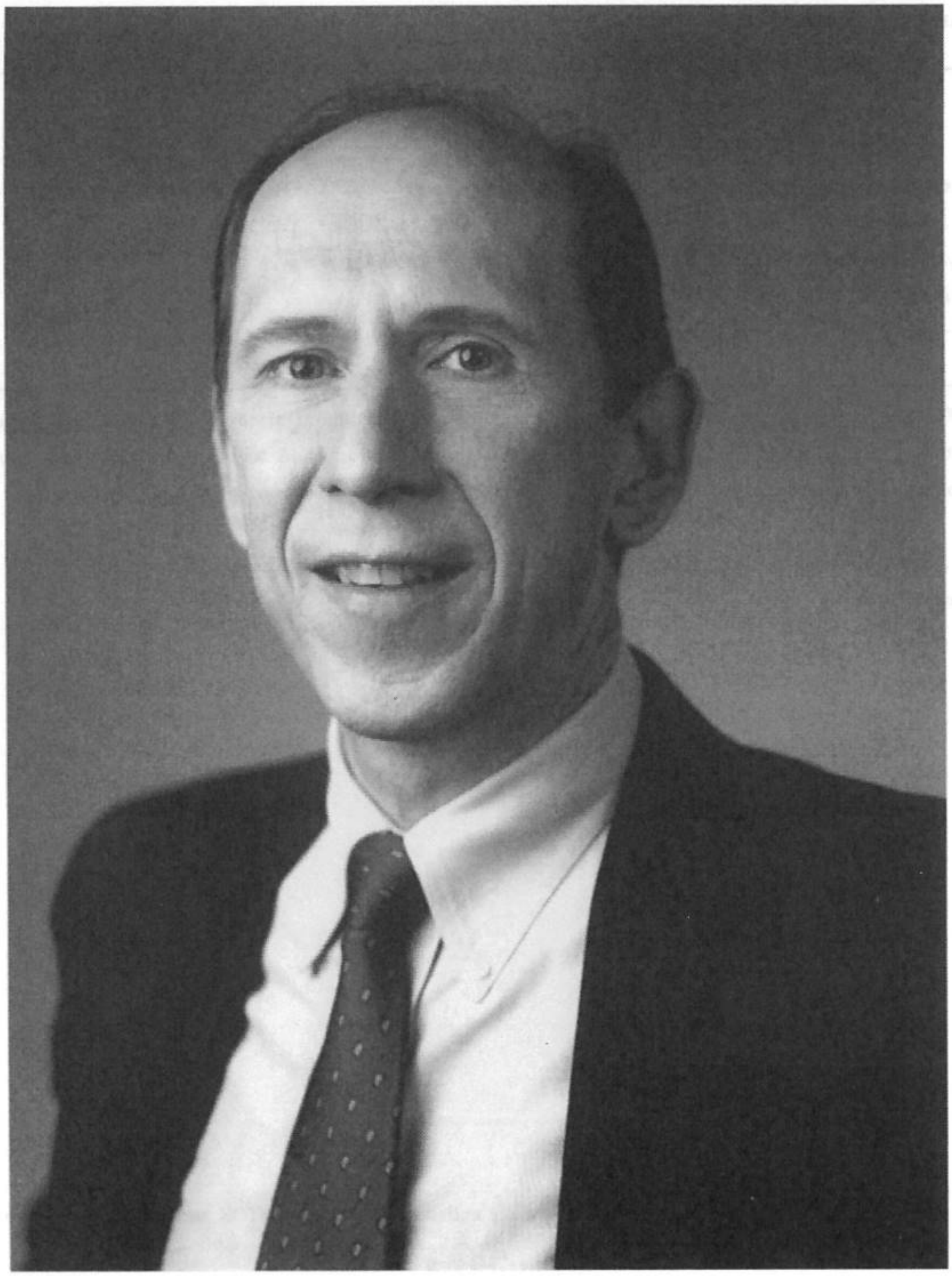
111. Louis D. Brandeis, Other People's Money and How the Bankers Use It 62 (National Home Library Foundation ed. 1933).

Notes & Observations

43

OMITTED

44



Kenneth R. French

THE JOURNAL OF FINANCE • VOL. LXIII, NO. 4 • AUGUST 2008

Presidential Address: The Cost of Active Investing

KENNETH R. FRENCH*

ABSTRACT

I compare the fees, expenses, and trading costs society pays to invest in the U.S. stock market with an estimate of what would be paid if everyone invested passively. Averaging over 1980–2006, I find investors spend 0.67% of the aggregate value of the market each year searching for superior returns. Society's capitalized cost of price discovery is at least 10% of the current market cap. Under reasonable assumptions, the typical investor would increase his average annual return by 67 basis points over the 1980–2006 period if he switched to a passive market portfolio.

HOW MUCH DO INVESTORS SPEND TRYING to beat the market? To answer this question, I start by estimating the total amount society spends to invest. I measure four components: the fees and expenses investors pay for mutual funds, including open-end funds, closed-end funds, and exchange-traded funds; the investment management costs of institutional investors; the fees investors pay for hedge funds and funds of hedge funds; and the costs all investors pay to trade. I then compare these costs to what society would pay if all investors held a passive market portfolio. The difference is the cost of active investing.

Consider a small but representative investor whose initial investment strategy is the value-weight combination of all investors' strategies. Because the value-weight combination of all investors' portfolios is the market portfolio, the representative investor's initial return is the gross return on the market minus the value-weight average of all investors' costs. How would his return be

*Kenneth R. French is at the Amos Tuck School of Business, Dartmouth College. I had an extraordinary amount of help on this project. Keith Ambachsteer, Van Anthony, Brad Asness, Cliff Asness, Paul Bennett, Jack Bogle, David Booth, Stephen Brown, Kent Clark, Sean Collins, Shrikant Dash, Bob Deere, Doug Diamond, Robert Dintzner, Jim Dunne, Scott Esser, Maribeth Farley, Beth French, Will Goetzmann, Henry Gray, John Griswold, David Hall, Frank Hatheway, Ken Heinz, Terry Hendershot, Sarah Holden, John Howard, Susan Hume McIntosh, Antti Ilmanen, John Liew, Ananth Madhavan, Bernie Madoff, Kevin Maloney, Dave Martin, Charles McNickle, Mike Mendelson, Nir Messafi, Mark Mitchell, Catherine Newell, Terry Odean, John Rekenthaler, Eduardo Repetto, Savina Rizova, Mark Rubinstein, Dave Schneider, Bob Shiller, Jeff Smith, George Sofianos, Hans Stoll, Dick Thaler, Ingrid Tierens, Paula Volent, and Wayne Wagner gave me valuable information and advice, and Robert Batt, Bob Burnham, Marianna Paskar, Savina Rizova, and Blake Tatsuta provided excellent research assistance. I have also benefited from extensive conversations with Brad Barber, Jon Lewellen, Hubert Lum, Sunil Wahal, and especially Gene Fama. I thank TD Ameritrade, CEM Benchmarking, the Common Fund, Dimensional Fund Advisors, Greenwich Associates, Hedge Fund Research, Morningstar, Standard and Poor's, and the Securities and Exchange Commission for data.

affected if he switched to a passive market portfolio? My answer depends on a key assumption: There is no net transfer between the passive market portfolio and other investors. The manager of the passive portfolio, for example, does not lose to or take advantage of other investors when he trades. With this assumption, which I support with empirical evidence below, the return on a passive market portfolio is the gross market return minus the cost of investing passively. Thus, a small representative investor who switches to a passive market portfolio increases his return by the difference between the value-weight average of all investors' costs and the cost of investing passively. Equivalently, his premium for switching is the difference, per dollar invested, between society's total cost of investing and the cost in the passive scenario. (The logic here is similar to the logic of Malkiel (1973), Sharpe (1991), Buffett (2006), and others.)

The no-net-transfer assumption guarantees that, in aggregate, the search for trading gains is doomed. Before considering costs, a trading gain for one active investor must be a loss for another. When we include their higher fees, expenses, and trading costs, it is clear that active investors are playing a negative sum game. This does not mean, however, that the cost of active investing is a pure loss to society. In aggregate, active investors almost certainly improve the accuracy of financial prices. This, in turn, improves society's allocation of resources. Thus, my estimate of the cost of active investing also measures society's cost of price discovery. I offer no evidence on whether society is buying too little or too much of this good. Price discovery, however, is an externality—each active investor pays the full cost of his efforts but captures only a tiny slice of the benefit—so there is no reason to think active investors purchase the optimal amount of price discovery.

I limit the scope of the paper by considering only the costs of investing in U.S. equity. Most of the results are for 1980–2006, but when they are available, I include data for 2007. The average of the annual estimates for 1980–2006 implies investors spend 0.67% of the value of all NYSE, Amex, and NASDAQ stocks each year trying to beat the market.[1] Under the no-net-transfer assumption, this means that an investor who holds a passive market portfolio outperforms the value-weight average of all active and passive investors by 67 basis points a year from 1980 to 2006.

If the expected real return on U.S. equity is roughly 6.7% and we assume the annual dollar cost of active investing will not increase in the future, society's capitalized cost of price discovery is about 10% of the current value of the market. Estimates of the equity risk premium in Fama and French (2002) and Graham and Harvey (2005), however, suggest that the expected real return on the market is substantially below 6.7%. If so, the capitalized cost of price discovery is above 10% of the current market cap. Moreover, the data imply that the cost of active investing will grow with the aggregate market cap. This expected growth pushes the capitalized cost even higher. Thus, 10% of the current value of the market is a conservative estimate of the capitalized cost of price discovery.

[1] Bogle (2008) offers a more inclusive estimate of society's cost of investing for 2007.

The first step in my analysis, in Section I, is to estimate the allocation of publicly traded U.S. equity among groups of investors. Direct holdings by individuals decline a lot over time. Individuals hold 47.9% of the market in 1980 and only 21.5% in 2007. This decline is matched by an increase in the holdings of open-end mutual funds, from 4.6% in 1980 to 32.4% in 2007. The shift from direct holdings to open-end funds has an important implication. Some argue that mistakes by retail investors are a reliable source of trading gains for other investors. If so, competition for these gains must be fierce later in the sample as an expanding group of professional investors fights for a shrinking pool of mistakes.

I examine the cost of mutual funds in Section II. Driven by a steady decline in the loads open-end fund investors pay, the fees and expenses for mutual funds fall from 2.08% of assets under management in 1980 to 0.95% in 2006. The investment management costs for institutions, which I estimate in Section III, are lower. Their value-weight average cost is only 34 basis points in 1980 and 23 basis points in 2006. Institutional costs decline over time for two reasons. First, the costs they pay for active and passive investments decline. Second, and more interesting, institutions shift a large portion of their U.S. equity holdings from active to passive over time.

In Section IV, I use data on individual hedge funds to estimate the fees clients pay to invest in U.S. equity-related funds. The average annual hedge fund fee for 1996–2007 is a hefty 4.26% of assets, and, because they pay two layers of fees, the average for clients who buy through funds of hedge funds is even higher, 6.52% per year. My analysis of trading costs, in Section V, follows Stoll (1993). I use data from the Securities and Exchange Commission (SEC) to measure the total commissions and market-making gains brokers and dealers earn by trading U.S. stocks.

My bottom line is in Section VI. I compare the resources investors actually spend in the U.S. market—the fees and expenses paid for mutual funds, the investment management costs paid by institutions, the fees paid to hedge funds and funds of funds, and the transaction costs paid by all traders—with what investors would spend if everyone followed a passive strategy. The difference between the actual and passive estimates is the cost of active investors' search for superior returns.

Standardized by the total value of NYSE, Amex, and NASDAQ stocks, the amount investors spend trying to beat the market is surprisingly stable; the cost is between 61 and 74 basis points in 24 of the 27 years from 1980 to 2006 and in every year after 1990. Although the total amount is relatively constant, the components change a lot over time. Because the amount invested in mutual funds increases so much through time, for example, the expenditures on fund fees and expenses increase from 0.11% of total market cap in 1980 to 0.32% in 2006. The fees for U.S. equity-related hedge fund investments also grow a lot, from essentially zero early in the period to 0.13% of the total value of U.S. equity in 2006. These increases are offset by a dramatic drop in the cost of trading. Despite a sharp increase in trading volume, the aggregate cost of trading U.S. equity falls from 0.55% of total market cap in 1980 to only 0.21% in 2006. Thus, measured relative to the value of U.S. equity, investors shift their expenditures from trading to asset management, but the total amount spent to beat the market is never far from the 1980–2006 average of 67 basis points.

My estimate of the resources consumed in the search for superior returns does not include several potentially important costs. Retail brokers, for example, borrow from their customers at below market rates and make margin loans to them at above market rates. Although the income from these activities is part of the revenue firms earn for trading—and part of their customers' cost of trading—I miss this in my estimate of the resources investors spend trying to beat the market. Fees for wealth management, such as financial and estate planning, are not a cost of active investing, but my estimate should include advisor fees that are for advice about undervalued stocks and winning investment strategies.

I intentionally omit transfers between investors. An active investor, for example, may pay a large market impact cost to trade quickly. If the counterparty is a broker, this trading cost is included in the market-making gains the broker reports to the SEC, and it is in my estimate of the resources society spends to trade. If the counterparty is another investor, however, the market impact cost is just a transfer, reducing one investor's return and increasing another's by the same dollar amount. Thus, it is not a cost to society. Similarly, to a taxable investor choosing between active and passive strategies, the extra tax burden that typically accompanies active trading is a cost. From society's perspective, however, extra taxes are just a transfer, so I do not include them in my estimate of the resources society spends to beat the market.

Most security lending payments are also transfers—one investor pays to borrow the security and the other receives the payment—so they are appropriately excluded from my estimate of the cost of active investing. The trading desk that arranges a security loan, however, typically retains part of the payment as compensation for its services and this does belong in my estimate. Similarly, the interest retail brokers earn lending securities held in street name is part of their compensation for providing trading services. The results below miss both of these costs.

I overstate the cost of active investing in at least two ways. First, the fees and expenses I measure include manager compensation. Many managers invest in their own funds, so my estimates include payments managers make to themselves. This is not much of a problem for mutual funds since managers own only a small fraction of aggregate fund assets, but it may be significant for hedge funds.

More important, I assume most investors switch to a market portfolio in the passive scenario. (Individuals with direct stock holdings and employee stock ownership plans continue to hold their actual portfolios.) There are several reasons, however, why passive investors might choose something other than a market portfolio. Taxable investors have an incentive to avoid realizing short-term gains and to defer long-term gains. Investors with specific social concerns might favor some securities over others. And, in the spirit of Merton (1973) or Ross (1976), some investors might shift away from the market portfolio because they prefer a different multifactor risk-return trade-off. To the extent that such deviations from the market portfolio increase the cost of investing in the passive scenario, I overstate the incremental cost of active investing.

Finally, I focus on the monetary cost of active investing, but most active investors bear a second cost: Their portfolios are not as well diversified as the market portfolio. The fact that the average investor could increase his return and lower his risk simply by switching to a passive market portfolio raises an obvious question. Why do active investors continue to play a negative sum game? I summarize the paper and address this question in Section VII. An extensive Appendix completes the paper.

I. Allocations

Table I describes the ownership of U.S. common equity from 1980 to 2007. Most of the information I use to measure these allocations is from the December 6, 2007 release of the Federal Reserve Board's Flow of Funds Accounts, which reports the total amount of corporate equity held by various investor groups. The adjustments I make to convert these estimates to the allocations in Table I are described in the Appendix.

There are several interesting patterns in the allocations in Table I. In 1980 individuals hold the biggest share of U.S. common equity, 47.9%. Direct holdings shrink to about 27% in 1994–1996, jump back to 36% in 1999, 2000, and 2001, and then fall steadily to only 21.5% at the end of October 2007. The growth in open-end mutual funds is equally dramatic, from 4.6% in 1980 to 32.4% in 2007, and although the yearly changes are not perfectly aligned, the total increase about matches the reduction in direct holdings.

The shift from direct holdings to open-end funds has at least two important implications. First, in the analysis below, only the fees and expenses for hedge funds are higher than those for open-end funds. Since I assume there are no fees or expenses associated with direct holdings, the shift to funds pushes up my estimate of society's cost of investing. But there is also a benefit. Most people who hold stocks directly are more poorly diversified. (See, for example, Barber and Odean (2000), or Goetzmann and Kumar (2008).) Thus, although the shift to open-end funds increases my estimate of society's overall cost of investing, it also reduces the typical investor's risk. Second, some claim that retail investors are a reliable source of trading gains for mutual funds, hedge funds, and other institutional investors. If so, the shift from direct holdings to open-end funds suggests these gains become scarcer later in the sample as an expanding pool of professional managers competes for a shrinking pool of retail mistakes.

The Fed's allocations include not only the U.S. equity I focus on, but also foreign equity owned by U.S. residents and institutions. Table I reports the value of these foreign holdings as a fraction of U.S. investors' total equity portfolio. Readers familiar with the literature seeking to explain why investors do not diversify internationally (e.g., French and Poterba (1991), Karolyi and Stulz (2003), and Ahearne, Griever, and Warnock (2004)) may be surprised that this

Table I

Allocation of Publicly Traded U.S. Common Equity, in Percent, 1980–2007

The table reports the percent of U.S. equity held by various investors. Most of the data are from the December 6, 2007 release of the Federal Reserve's Flow of Funds Accounts. The steps used to compute the allocations are described in the Appendix. Direct holdings is the fraction of U.S. equity held by individuals. The allocation to closed-end funds, CEFs, includes unit investment trusts. The holdings of foreign investors are included in the other categories and foreign holdings is the fraction of U.S. investors' equity portfolio in foreign stock.

	Direct Holdings	Open-end Funds	CEFs	ETFs	DB Plans	DC Plans	ESOPs	Public Funds	Nonprofits	Banks and Insurance	Hedge Funds	Foreign Investors	Foreign Holdings
1980	47.9	4.6	0.5	0.0	18.1	3.9	2.8	4.4	8.3	9.4	0.0	7.6	2.0
1981	45.9	4.4	0.5	0.0	19.0	3.7	3.5	5.1	7.9	10.1	0.0	8.1	1.9
1982	42.4	5.0	0.4	0.0	21.1	3.5	4.6	5.5	7.2	10.2	0.0	8.2	1.7
1983	39.5	6.3	0.4	0.0	21.4	3.4	5.0	6.7	6.7	10.5	0.0	8.4	2.2
1984	37.3	7.0	0.3	0.0	21.8	3.0	6.7	7.4	6.3	10.2	0.0	8.4	2.2
1985	35.4	7.6	0.3	0.0	22.5	3.3	7.7	7.2	6.0	10.1	0.0	8.4	2.9
1986	37.4	9.4	0.4	0.0	20.3	2.6	6.6	7.6	6.2	9.4	0.0	9.7	4.1
1987	36.1	10.4	0.6	0.0	18.4	3.9	6.9	8.5	6.0	9.3	0.0	9.9	5.2
1988	39.3	9.6	0.7	0.0	15.5	3.8	6.1	9.5	6.5	8.9	0.0	10.2	6.4
1989	38.3	10.3	0.7	0.0	14.8	3.7	6.6	9.9	7.2	8.5	0.0	10.6	7.8
1990	35.4	10.5	0.9	0.0	15.3	4.3	6.4	11.1	7.6	8.2	0.3	10.1	8.5
1991	35.0	10.4	0.9	0.0	15.5	4.0	6.9	11.8	6.3	8.7	0.4	9.3	9.1
1992	33.0	12.4	0.9	0.0	14.9	3.8	7.4	11.9	6.5	8.6	0.5	9.0	9.4
1993	29.7	15.7	0.9	0.0	14.4	4.1	7.4	11.8	6.1	9.1	0.8	8.6	13.7
1994	26.8	17.9	1.0	0.0	14.1	4.3	6.9	11.9	6.8	9.5	0.8	8.9	15.4
1995	26.7	19.6	1.1	0.0	13.2	4.1	6.3	12.3	6.6	9.5	0.7	9.2	14.7
1996	27.2	22.2	1.1	0.0	11.5	3.9	5.4	12.1	6.5	9.3	0.8	8.8	14.8
1997	29.5	23.4	1.0	0.1	9.8	3.9	4.4	11.7	6.1	9.3	0.8	9.4	13.6
1998	30.2	24.3	1.0	0.2	9.2	4.3	4.0	11.1	5.8	9.4	0.7	10.4	13.9
1999	36.0	24.7	0.8	0.2	7.4	3.5	3.1	9.9	4.9	8.9	0.6	10.3	14.3
2000	36.2	24.4	0.6	0.5	8.1	3.6	2.6	9.5	4.6	9.1	0.7	10.8	13.7
2001	36.0	23.6	0.5	0.7	8.8	3.5	3.0	10.0	4.0	9.2	0.8	11.4	13.1
2002	32.1	23.7	0.5	1.1	9.9	3.6	3.0	10.8	3.6	10.5	1.3	12.3	14.5
2003	29.9	25.5	0.6	1.2	9.8	3.7	3.2	11.0	3.4	10.4	1.3	13.2	16.8
2004	27.1	27.6	0.8	1.6	9.7	3.9	3.1	11.0	3.1	10.7	1.4	13.6	18.6
2005	26.1	28.8	0.9	2.0	9.1	3.9	3.0	10.9	2.9	10.8	1.5	14.0	22.3
2006	24.2	30.5	1.0	2.5	8.6	4.0	2.8	10.7	2.6	11.2	1.9	15.1	25.3
2007	21.5	32.4	1.1	3.0	8.5	3.8	2.8	10.6	2.3	11.8	2.2	16.3	27.2

fraction grows from 2.0% in 1980 and 8.5% in 1990 to a substantial 27.2% in 2007. Thus, in 2007 more than one-fourth of the average U.S. investor's equity portfolio is in foreign stocks.

Table I shows a fairly steady decline in the share of U.S. equity owned by foundations, endowments, and other nonprofits, from 8.3% in 1980 to 6.0% in 1985 and 2.3% in 2007. One might be tempted to attribute the decline to the well-known shift by endowments toward alternative investments. However, Greenwich Associates, a consulting and research firm, reports that the fraction of endowment assets invested in domestic equity drops by less than one-third between 1985 and 2006, from 47.4% to 34.2%, so this is not the full explanation. Part of the decline may be an artifact of the process I use to disentangle allocations to nonprofits and households in the Flow of Funds Accounts (described in the Appendix). If so, the decline in the direct holdings of individuals is even bigger than the estimates in Table I imply.

The allocation to defined contribution (DC) plans in Table I remains close to 4% throughout the period, but this is a bit misleading. To avoid double counting, the allocations to DC and defined benefit (DB) plans in Table I do not include the mutual funds they own. The omission has only a modest impact on the estimates for DB plans, but it has a big impact on the estimates for DC plans. Supplemental data in the Flow of Funds Accounts imply that the mutual fund holdings of DC plans grow from 0.3% of the value of the U.S. market in 1985 to a substantial 8.5% in 2006. Although these estimates include fixed income and foreign equity funds, it is clear that by 2006 DC plans own much more U.S. equity than the 3.8% allocation in Table I suggests.

Finally, in terms of their net holdings of U.S. equity, hedge funds are relatively unimportant. They grow from 0.3% of U.S. equity in 1990 to 2.2% in 2007. But we shall see that hedge funds play a big role when we look at costs.

II. Average Fees and Expenses for Mutual Funds

My estimates of the resources spent trying to beat the market combine the allocations to various groups, in Table I, with estimates of each group's cost of investing. To be conservative, I assume the only expenses individuals incur when they hold shares directly are trading costs, which are included in the aggregate estimates below. I ignore, for example, the time they spend managing their portfolios and the cost of subscriptions to Value Line and Morningstar. Similarly, I assume employee stock ownership plans (ESOPs) have no investment management costs.

I use reported expense ratios, from the mutual fund database maintained by the Center for Research in Security Prices (CRSP), and estimates of annuitized loads, from the Investment Company Institute (ICI), to measure the cost of investing in open-end funds. The average expense ratios in Table II weight funds by their assets under management at the beginning of the year, and include only those that invest predominantly in U.S. common equity. (The Appendix describes the steps used to identify U.S. equity funds. Fama and French (2008) analyze the returns on this set of funds.)

Table II

Fees and Expenses for Mutual Funds, in Basis Points, 1980–2006

The expense ratio for open-end mutual funds is the value-weight average of the reported values for U.S. equity funds in the CRSP mutual fund database. The annuitized load is from the Investment Company Institute and measures the value-weight average load paid by investors in equity funds. Total is the sum of the open-end expense ratio and annuitized load. Percent passive is also from the ICI and measures the fraction of U.S. equity fund assets invested in index funds. The value-weight average expense ratios for U.S. equity closed-end funds (CEFs) and U.S. equity exchange-traded funds (ETFs) are estimated using data from Morningstar.

	Open-end Mutual Funds				Expense Ratio	
	Expense Ratio	Annuitized Load	Total	Percent Passive	CEFs	ETFs
1980	70	149	219			
1981	71	167	237			
1982	75	128	203			
1983	76	113	190			
1984	82	114	196	1.0		
1985	80	105	185	1.1		
1986	81	101	183	0.8		
1987	86	96	182	0.9		
1988	96	97	193	1.2		
1989	94	84	178	1.6		
1990	93	76	169	2.3		
1991	90	65	155	2.9		
1992	96	59	155	3.6		
1993	96	50	146	3.9		
1994	98	47	145	3.9		
1995	96	42	139	4.7		
1996	93	40	134	5.8		
1997	92	35	126	7.3		
1998	90	30	120	9.0		
1999	91	27	117	9.7		
2000	96	24	119	9.8	96	
2001	97	19	116	10.9	92	20
2002	98	18	116	12.4	101	17
2003	96	17	113	12.4	98	18
2004	91	18	108	12.7	104	19
2005	87	16	103	12.5	103	20
2006	85	15	100	12.6	109	21

The value-weight average expense ratio for open-end funds grows from 70 basis points in 1980 to 96 basis points in 1988. It remains in a narrow band over the next 14 years and then falls from 98 basis points in 2002 to 85 in 2006. One might suspect that the decline in the average expense ratio at the end of the period reflects a shift from active open-end funds to lower priced passive funds. Table II shows that there is a shift to passive funds, from 1.0% of fund assets in 1984 to 12.4% in 2002, but it occurs before the average expense ratio falls. The growth of exchange-traded funds (ETFs) and competitive pressure from passive open-end funds, however, probably contribute to the decline.

The behavior of the average annuitized load in Table II is striking. It falls almost monotonically from 149 basis points in 1980 to only 15 basis points in 2006. (Barber, Odean, and Zheng (2005) make a similar point.) This drop, which is driven mostly by a shift toward no-load funds, has a big impact on the total fees and expenses paid by investors. The annual costs of open-end funds shrink from 2.19% of assets under management at the beginning of the period to 1.00% at the end.

Because closed-end funds and ETFs trade on exchanges, customers pay brokerage commissions rather than loads when they buy and sell these funds. The commissions are part of the aggregate trading costs measured below. Thus, I include only expense ratios in the fees and expenses for investments in U.S. equity closed-end funds and ETFs. The data I have on these funds, from Morningstar, are not as complete as those for open-end funds; I can compute annual value-weight average expense ratios for closed-end funds only from 2000 to 2006 and for ETFs from 2001 to 2006. I use the averages of these annual estimates before 2000 and 2001. Fortunately, ETFs are 0.5% or less of U.S. equity before 2001 and the allocation to closed-end funds never exceeds 1.1%, so imprecise estimates of the annual average expense ratios have little effect on my results. The average of the annual estimates for U.S. equity closed-end funds in Table II, 1.01%, is a bit higher than the average expense ratio for open-end funds over the same period, 0.93%. The 2001–2006 average for ETFs is only 19 basis points, which is not surprising given that most ETFs are variants of passive funds in that period.

III. Institutional Costs

The information I use to measure the investment expenses of institutional investors comes from two sources. CEM Benchmarking, Inc., a Toronto-based firm that monitors the investment activities of pension plan sponsors, provided annual estimates of the costs incurred by DB and DC plans when they make active and passive investments in the U.S. stock market. I combine these with estimates of the active and passive U.S. equity allocations of institutional investors from Greenwich Associates.

The Greenwich estimates are from surveys of DB plans, DC plans, public funds, and nonprofits, which include foundations and endowments through 1999 and only endowments thereafter. Greenwich has conducted surveys annually since 1986 and the respondents control a substantial portion of all institutional investments. For example, 1,950 institutions with more than six trillion dollars participated in the 2006 survey.

The results of the Greenwich surveys are in Table III. All four groups of institutions increase their allocation to passive over time. DB plans show the smallest increase, from 21.1% in 1986 to 31.2% in 2006. Nonprofits start with a meager 2.8% of their U.S. equity holdings invested passively, but finish with 28.7%. Public funds have the highest passive allocation throughout the period, with 25.8% in 1986 and a substantial 52.7% in 2006.

Table III

Percent of U.S. Equity Investments Allocated to Passive Strategies by Institutions and Investment Management Costs Incurred by Institutions, in Basis Points, 1986–2006

The percent invested passively, from Greenwich Associates, is the value-weight average fraction of their U.S. equity investments institutions allocate to passive strategies. The four institutional groups are defined benefit (DB) plans, defined contribution (DC) plans, public funds, and nonprofits, which include foundations and endowments through 1999 and only endowments thereafter. The passive and active investment management costs for DB plans are value-weight averages, from CEM benchmarking. The investment management costs for DB plans, public funds, and nonprofits are weighted averages of the passive and active DB costs. The investment management costs for DC plans use the annual passive and active DB costs plus the average annual difference between DC and DB costs. The average difference is 3.4 basis points for passive and 18.2 basis points for active. I use the 1991 estimates of passive and active costs for 1986–1990.

	Percent Invested Passively				Investment Management Cost					
					DB Plans					
	DB Plans	DC Plans	Public	Nonprofits	Passive	Active	DB Plans	DC Plans	Public	Nonprofits
1986	21.1	17.9	25.8	2.8			34	50	32	39
1987	24.6	26.2	29.0	9.6			32	46	31	37
1988	22.3	29.6	39.0	9.2			33	45	28	37
1989	25.6	29.7	46.0	12.7			31	45	25	36
1990	28.5	29.4	43.1	12.5			31	43	26	36
1991	32.5	31.9	44.6	14.3	7.9	40.4	30	43	28	36
1992	26.8	35.0	39.2	11.1	5.9	42.1	32	42	28	37
1993	25.6	33.8	46.4	17.0	6.7	43.1	34	44	26	37
1994	27.8	31.7	43.4	19.0	7.4	45.4	35	47	29	38
1995	27.5	32.2	40.0	23.8	6.0	43.4	33	45	28	34
1996	30.7	32.1	48.5	23.2	5.4	37.9	28	41	22	30
1997	29.5	33.7	52.2	18.1	4.9	36.3	27	39	20	31
1998	27.1	30.6	52.9	17.9	4.6	34.2	26	39	19	29
1999	30.0	34.0	54.2	20.2	3.8	34.0	25	37	18	28
2000	29.4	35.1	57.1	20.7	4.3	35.6	26	38	18	29
2001	30.5	32.5	51.9	22.0	4.5	37.2	27	40	20	30
2002	30.4	35.0	52.4	23.5	4.2	41.3	30	41	22	33
2003	32.7	34.6	55.2	36.4	2.8	37.8	26	39	18	25
2004	34.4	33.2	53.6	29.4	2.6	35.8	24	38	18	26
2005	31.2	34.6	53.7	25.8	2.7	37.0	26	38	19	28
2006	31.2	35.7	52.7	28.7	2.9	36.0	26	37	19	27

CEM Benchmarking's estimates of the cost of active and passive investing are based on a smaller sample of institutions. In 2006, for example, CEM has information on 141 DB plans and 99 DC plans. CEM tends to focus on larger plans, however, so those in the 2006 sample have 2.78 trillion dollars in total assets, with more than one trillion invested in publicly traded U.S. equity. The underrepresentation of smaller institutions probably has little impact on my estimates of the cost of active investing. First, because they have more assets to invest, larger institutions are more important for the aggregate values I am trying to measure. Second, estimates of what society spends to beat the market depend on the difference between the costs of active and passive investing. CEM's emphasis on large plans may reduce my overall estimates of the institutional cost of investing, but because economies of scale affect both active and passive costs, it has less effect on the difference.

CEM provides annual value-weight averages of the costs incurred by DB and DC pension plans for passive and active investments in U.S. common equity. The costs include auditing, consulting, oversight, and custodial charges, compensation and other employee costs, and investment management fees for externally managed strategies. The estimates for DB plans, which are available for 1991–2006, are in Table III. As expected, active strategies cost a lot more than passive strategies. The average of the annual estimates for active, 38.6 basis points, is eight times the average for passive, 4.8 basis points. Both passive and active costs decline over time. The average cost for active strategies in DB plans falls from 40.4 basis points in 1991 to 36.0 basis points in 2006, and the average cost for passive strategies falls from 7.9 basis points to only 2.9 basis points. The decline in costs is not caused by a change in the DB plans sampled. Similar declines are observed if the sample is limited to only plans with data for the whole 16-year period.

My annual estimates of the costs paid by (i) DB plans, (ii) public plans and state and local governments, and (iii) foundations, endowments, and other nonprofits, in Table III, combine the average costs of passive and active strategies in DB plans from CEM with the allocations between passive and active from Greenwich. Specifically, the investment management cost for a group is the passive cost for DB plans times the group's allocation to passive strategies plus the active cost times the group's allocation to active strategies. Since the CEM data are not available before 1991, I use the 1991 estimates of the cost of active and passive strategies for 1986–1990.

The CEM data for DC plans do not start until 1997. Perhaps because the sample of DC plans is smaller than the sample of DB plans, the annual cost estimates for DC plans (not reported) are more volatile than the estimates for DB plans. Because of this volatility, I use the annual DB cost plus the average difference between the costs for DC and DB plans for 1997–2006 to measure the annual cost of active and passive DC strategies. The investment costs for DC plans are generally higher than the costs for DB plans. The average difference is 3.4 basis points for passive strategies and 18.2 basis points for active strategies.

The estimated costs for all four institutional groups in Table III decline between 1986 and 2006. The smallest drop is for DB plans, from 34 to 26 basis points. The cost for each of the other three groups declines by 12 or 13 basis points—from 50 to 37 basis points for DC plans, from 32 to 19 basis points for public funds, and from 39 to 27 basis points for nonprofits. These reductions are the result of the decline in the costs of active and passive strategies and, more important, the shift over time from active to passive investments. This shift toward passive strategies is in sharp contrast to the contemporaneous growth of hedge funds, examined next.

IV. Hedge Fund Fees

Hedge fund fees often have two components. A fee of "2 and 20," for example, means that investors pay an annual management fee of 2% of the assets in the fund plus a performance fee of 20% of profits. Because of the performance fee, the aggregate compensation paid to hedge fund managers each year depends on the specific return earned by each fund in the industry. I use data from Hedge Fund Research (HFR) to estimate the fees on individual hedge funds and funds of hedge funds from May 1996 to December 2007.

Hedge funds trade stocks, bonds, currencies, and other securities in markets around the world. Since I am measuring the resources spent trying to beat the U.S. stock market, I have to estimate the fraction of hedge fund assets used in U.S. equity strategies. HFR assigns hedge funds to one of several categories, such as merger arbitrage, event driven, and sector funds. I use their categories to sort funds into three groups. I assume funds in the first group use 100% of their assets for equity strategies, those in the second use 50%, and those in the third do not use any of their assets for equity trading. I then use the weight of the U.S. in the world equity portfolio to estimate the fraction of equity-related assets used for trading in the U.S. (The Appendix describes this process in detail.)

Table IV shows HFR's annual estimates of the assets invested in the hedge fund industry and my estimates of the assets in U.S. equity-related strategies. Total hedge fund assets grow rapidly during the sample period, from less than 40 billion dollars in 1991 to 185.8 billion in 1996 and 1,464.5 billion at the beginning of 2007. Investment in U.S. equity-related strategies keeps pace with the total; at the beginning of 2007 there are 458.6 billion dollars in these strategies.[2]

A large fraction of hedge fund assets is held by funds of funds. In 2007, for example, 655.9 billion dollars—about 45% of all hedge fund assets—are invested in funds of funds. The HFR database puts all funds of funds in the same category, so I am unable to isolate those that focus on U.S. equity-related strategies. In the analysis below I assume that funds of funds are invested proportionately between hedge funds that are related to U.S. equity and those that are not.

Table IV also reports annual value-weight averages of the fees for funds of funds and U.S. equity-related hedge funds for 1996–2007. Quoted hedge fund fees increase over the sample period. The value-weight average management fee rises from 0.92% in 1996 to 1.28% in 2007, and the average quoted performance fee rises from 18.24% to 19.15% over the same period. There is less variation in the average management fee for funds of funds, but their average quoted performance fee declines a lot over time, from 9.45% in 1996 and 11.41% in 1998 to 6.94% in 2007.

[2] Because hedge funds use leverage and take long and short positions, the total assets in U.S. equity-related strategies, in Table IV, differ from the net holdings of U.S. equity implied by the allocations in Table I. The Appendix explains how I calculate the estimates in both tables.

Table IV

Assets Invested in Hedge Funds and Funds of Hedge Funds, 1991–2007, and Hedge Fund and Fund of Fund Fees, 1996–2007

Assets invested are in billions of dollars and are measured at the beginning of the year. The total for all hedge funds is from Hedge Fund Research (HFR). I use HFR's categorization of funds and the U.S. weight among all developed equity markets to estimate the U.S. equity-related hedge fund assets. The estimate of U.S. equity-related fund of fund assets assumes funds of funds invest proportionately among hedge fund categories. All but the last column of fees are value-weight averages of individual fund fees computed using the HFR database and are in percent. The last column is in billions of dollars. The management fee (Mgmt Fee) is a fixed percent of assets in the fund. The quoted performance fee is a fraction of the fund's profits. The actual performance fee is measured relative to the beginning-of-year assets. The Appendix describes how the actual performance fee is computed. Averages of the annual fees (and standard errors) are reported for 1996–2007 and 2000–2007. The data for 1996 start in May. The management fee and quoted performance fee for 1996 are annualized, but the actual performance fee is not.

	Assets Invested			Fees									
		U.S. Equity-related		U.S. Equity-related Hedge Funds				Funds of Hedge Funds				Hedge Fund plus Fund of Fund	
					Performance Fee				Performance Fee				
	All Hedge Funds	Hedge Funds	Funds of Funds	Mgmt Fee	Quoted	Actual	Total	Mgmt Fee	Quoted	Actual	Total	Percent	Dollars
1991	38.9	8.3	0.4										
1992	58.4	14.1	4.8										
1993	95.7	24.9	9.6										
1994	167.8	39.8	17.9										
1995	167.4	38.6	17.2										
1996	185.8	46.2	14.0	0.92	18.24	3.79	4.71	1.16	9.45	3.41	4.57	9.27	2.8
1997	256.7	72.4	14.8	1.05	18.40	5.40	6.45	1.28	9.29	2.57	3.85	10.30	5.2
1998	367.6	126.5	25.5	0.98	18.25	2.56	3.54	1.18	11.41	0.21	1.39	4.93	4.8
1999	374.8	141.4	28.6	1.03	18.24	5.91	6.94	1.33	8.24	1.68	3.01	9.95	10.7
2000	456.4	176.1	29.4	1.09	18.42	2.35	3.44	1.26	7.99	0.38	1.64	5.08	6.5
2001	490.6	203.3	34.6	1.20	18.93	1.51	2.71	1.24	7.22	0.27	1.51	4.22	6.0
2002	539.1	243.4	46.3	1.24	19.09	1.38	2.62	1.11	7.12	0.34	1.45	4.07	7.0
2003	625.6	253.1	83.7	1.25	19.12	3.40	4.65	1.14	6.96	1.23	2.37	7.02	13.8
2004	820.0	313.4	112.1	1.23	19.03	2.28	3.51	1.12	6.98	0.78	1.90	5.41	13.1
2005	972.6	348.1	128.4	1.26	19.03	2.18	3.44	1.15	7.09	0.47	1.62	5.06	14.1
2006	1105.4	372.8	133.1	1.27	18.95	3.25	4.52	1.17	7.12	0.76	1.93	6.45	19.4
2007	1464.5	458.6	205.4	1.28	19.15	3.35	4.63	1.12	6.94	0.73	1.85	6.48	25.0
1996–2007				1.16 (0.03)	18.74 (0.11)	3.11 (0.41)	4.26 (0.39)	1.20 (0.02)	7.98 (0.41)	1.07 (0.29)	2.26 (0.30)	6.52 (0.63)	
2000–2007				1.23 (0.02)	18.97 (0.08)	2.46 (0.28)	3.69 (0.29)	1.16 (0.02)	7.18 (0.12)	0.62 (0.11)	1.78 (0.10)	5.47 (0.38)	

The often mentioned "2 and 20" overstates the typical hedge fund fee. In 2007, for example, the value-weight average management fee is 1.28%, not 2%, and the value-weight average quoted performance fee is 19.15%, not 20%. (These averages do not change much if I use all funds, not just U.S. equity-related assets.) The standard "1 and 10" is a better description of the average management fee for funds of funds, but it overstates the average quoted performance fee by about 3% after 2001.

How much do hedge fund investors pay to beat the market? Averaging over the annual value-weight averages for 1996–2007, the typical investor in U.S. equity-related hedge funds pays a management fee of 1.16% of assets and a realized performance fee of 3.11%. Equivalently, the hedge fund industry must generate average annual abnormal returns of 4.26% over this period for the typical investor to break even. The average performance fee is extraordinarily high in 2 of the first 4 years of the sample, 5.40% in 1997 and 5.91% in 1999. If we focus on the results for 2000–2007, the average drops a bit, but investors still pay an annual combined fee of 3.69% over this 8-year period.

Hedge fund clients who invest through funds of funds pay two layers of fees. Averaging over the annual averages for 1996–2007, the typical fund of fund investor pays 2.26% in fund of fund fees and 4.26% in hedge fund fees. Thus, the underlying hedge funds must generate an average abnormal annual return of 6.52% for him to break even. If we throw out the first 4 years, the averages are lower—1.78% for the fund of fund fee and 5.47% for the total fee—but the threshold for investor success is still high.[3]

These estimates include only hedge fund and fund of fund fees. Among other things, they ignore the legal expenses, accounting and auditing fees, custody costs, and board fees that are paid by the funds. Although I am not able to measure these costs for hedge funds, I can use data from CRSP to infer the cost of comparable services for mutual funds. Specifically, the cost is the difference between a mutual fund's reported expense ratio and the sum of its management and 12b-1 fees, which are both available in the CRSP database after 2000. The average of the annual value-weight averages for U.S. equity mutual funds for 2001 to 2006 is 21 basis points.

My estimates of hedge fund costs also miss most of the payments they make to their prime brokers. These include financing costs, security lending fees, and charges for settling transactions done at other brokers. I do, however, capture the trading costs of hedge funds in the estimates I discuss next.

V. Trading Costs

Stoll (1993) develops a simple way to measure the aggregate cost of trading. The total commissions, bid-ask spreads, and other costs investors pay for trading services must equal the total revenue brokers and dealers receive for those

[3] Brown, Goetzmann, and Liang (2004) use the TASS hedge fund database to estimate realized performance fees for 1995–2003, and their annual average is higher than mine in 6 of the 8 years our periods overlap.

services. As Stoll (1993) shows, one can measure this revenue with information from the Financial and Operational Combined Uniform Single (FOCUS) reports that registered securities firms must file with the Securities and Exchange Commission each year. The trading revenue in the FOCUS reports includes commissions, which firms earn when they facilitate agency trades as a broker, and the gains or losses firms earn from market making. The process I use to extract this information, which is described in the Appendix, is almost identical to that used in Stoll (1993).

The FOCUS reports do not allow me to estimate three important components of trading revenue. Firms trading for retail investors are able to borrow money from clients at below market rates (typically through cash sweep accounts), make margin loans to clients at above market rates, and earn revenue by lending securities held in street name, including those in margin accounts.

Consider Charles Schwab, a large discount brokerage firm. The firm's financial statements show that in 2006 Schwab brokerage clients had an average daily balance of 17.86 billion dollars in interest-bearing cash accounts, with an average return of 2.38%. At the same time, Schwab lent clients 10.25 billion in margin loans at an average rate of 8.17%. As a rough estimate, the 5.79% spread in interest implies Schwab added 590 million dollars to its 2006 revenue by borrowing 10.25 billion dollars from some clients and lending it to others. And that still leaves 7.61 billion in the cash accounts. If Schwab invested this money in 30-day Treasury bills, which returned 4.81% in 2006, the opportunity to borrow 7.61 billion at 2.38% added another 185 million to its income. The total revenue Schwab earned by borrowing from and lending to its brokerage clients in 2006, 775 million dollars, almost matches the 785 million it reported in commissions and trading gains for the year.

Of course, this revenue is not free. In a competitive market, it is simply part of the compensation Schwab and other firms receive for providing brokerage services. This revenue and the revenue retail brokers earn by lending securities held in street name belong in my estimates of the total cost of trading. Unfortunately, I cannot isolate this income in the FOCUS data and few firms provide Schwab's level of detail in their financial statements. As a result, the revenue is missing from my estimates of trading costs.

Before turning to the estimates of cost, it is useful to look at the amount of trading investors do each year. Figure 1 shows the annual turnover of U.S. stocks from 1926 to 2007. The estimates use data from CRSP and include NYSE, Amex (starting in July 1962), and NASDAQ (1973) stocks with share codes of 10 or 11. (The Appendix explains how I deal with the double counting of trades on NASDAQ.) The turnover for a year is the sum of the 12 monthly estimates, which I measure as the ratio of the total dollar volume for the month (shares traded times beginning-of-month price) divided by the total market cap at the beginning of the month.

The general pattern in Figure 1 is striking. Turnover is above 110% in the 1920s. It reaches a high of 143% in 1928, then plunges with the market to 52% in 1932. By 1938 it is below 20%. In light of recent experience, it is perhaps surprising that annual turnover remains close to or below 20% from 1938 to

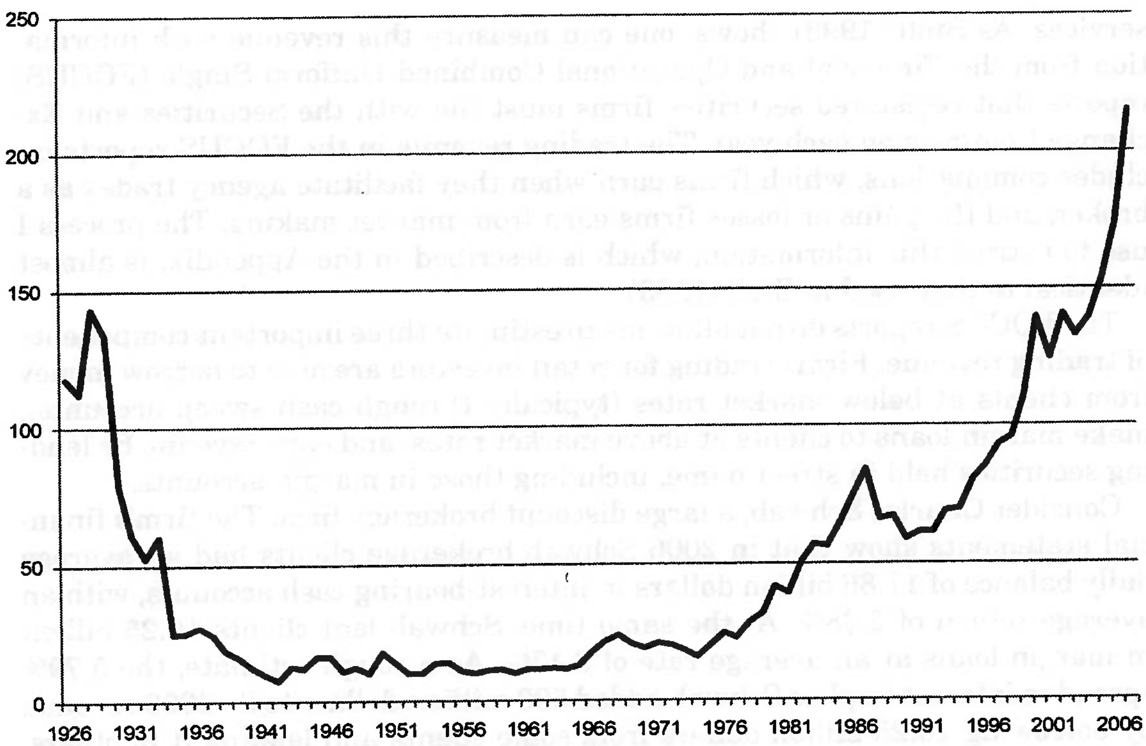


Figure 1. Annual turnover of NYSE, Amex, and NASDAQ stocks, in percent, 1926–2007.

1975. Turnover rises fairly steadily over the next three decades from 20% in 1975 and 59% in 1990, to an impressive 173% in 2006 and 215% in 2007.

Because they are not operating companies with CRSP share codes of 10 or 11, ETFs are not in the turnover in Figure 1. During the last few years of the sample, however, ETFs are heavily traded. Standard & Poor's Depository Receipts (Spiders) are the most extreme, with total volume of 2.3 billion dollars in 2006 and 5.9 billion in 2007. If I include domestic equity ETFs in my measure, aggregate turnover jumps from 173% to 208% in 2006 and from 215% to 284% in 2007.

What explains the extraordinary growth in trading between 1975 and 2007? Reduced costs are surely part of the story. The introduction of negotiated brokerage commissions in 1975, the development of electronic trading networks, the decimalization of stock prices in 2000 and 2001, and the SEC's implementation of rules designed to increase market transparency and liquidity, such as Reg NMS, all reduce the cost of trading U.S. equities during this period. But even at the end of the sample, trading is not free. From the perspective of the negative sum game, it is hard to understand why equity investors pay to turn their aggregate portfolio over more than two times in 2007.

The estimates from the FOCUS data, in Table V, confirm that the cost of trading falls a lot between 1980 and 2006. In fact, despite the explosive growth of trading during the last 6 years of the period, the *total* amount investors pay to trade declines by more than 35%, from 50.7 billion dollars in 2000 to 32.1 billion in 2006. The decline in the cost of trading is even more striking if we

Table V

Annual Revenue Received by Securities Firms for Executing Trades of U.S. Equity, 1980–2006

The data come from the FOCUS reports that broker and dealers file annually with the SEC. Commissions from exchange trades and OTC trades are commissions received for executing trades on an exchange and over the counter. Gains from market making include trading profits from OTC equities, gains on derivative trading desks in equity products, and gains on firm security trading accounts with associated hedges. Total revenue and the three components of total revenue are measured in billions of dollars. The Appendix describes how these values are calculated. Cost relative to volume, in basis points, is total revenue divided by total dollars traded on the NYSE, Amex, and NASDAQ.

	Commissions from				
	Exchange Trades	OTC Trades	Gains from Market Making	Total Revenue	Cost Relative to Volume
1980	4.1	1.0	1.0	6.1	146
1981	3.8	1.1	0.8	5.8	131
1982	4.3	1.2	0.9	6.4	117
1983	5.9	2.0	1.9	9.8	106
1984	5.0	1.7	1.3	8.0	88
1985	5.8	2.2	1.8	9.7	82
1986	7.2	2.8	2.5	12.5	73
1987	8.7	3.2	2.6	14.5	63
1988	6.2	2.5	2.0	10.7	66
1989	7.2	2.6	2.2	12.1	64
1990	6.1	2.6	2.0	10.7	64
1991	7.2	3.3	3.2	13.8	71
1992	7.8	4.1	3.9	15.8	69
1993	9.1	5.3	4.8	19.2	64
1994	8.6	5.2	4.5	18.3	57
1995	10.3	6.5	5.5	22.2	51
1996	11.3	8.3	6.6	26.2	45
1997	13.2	9.5	7.2	29.8	37
1998	14.6	10.4	7.7	32.7	32
1999	16.3	14.2	9.3	39.8	27
2000	18.2	17.6	14.9	50.7	23
2001	16.3	12.7	8.0	36.9	23
2002	16.0	12.8	4.6	33.4	21
2003	14.0	13.8	3.8	31.7	22
2004	13.8	15.0	3.6	32.4	16
2005	13.0	13.9	3.9	30.7	13
2006	12.2	13.7	6.2	32.1	11

standardize by the amount traded. Measured relative to total volume, the cost of trading declines (or remains constant) in all but 3 years between 1980 and 2006. The cumulative effect is a 92% reduction in trading costs, from 146 basis points in 1980 to a tiny 11 basis points in 2006. As we see next, this reduction has a significant effect on the resources investors spend in their search for superior returns.

VI. The Cost of Trying to Beat the Market

Table VI summarizes my estimates of the amount society pays to invest in the U.S. stock market. There are four components: the fees and expenses paid by those who purchase open-end funds, closed-end funds, and exchange-traded funds; investment management costs paid by institutions; fees paid by hedge fund investors; and trading costs paid by all investors. To make the costs easier to interpret, I standardize each year's dollar cost by the average capitalization

Table VI
Society's Standardized Cost of Investing, in Basis Points, 1980–2006

The standardized cost is the total dollar cost of investing divided by the aggregate market cap, which is the average of the 12 beginning-of-month values of all NYSE, Amex, and NASDAQ stocks with CRSP share codes of 10 or 11. The aggregate market cap is in billions of dollars. The contribution of mutual funds to the standardized cost is the total percent of U.S. equity in funds, from Table I, (Allocation, in percent) times the value-weight average of the fees and expenses of mutual funds, in Table II (Fees, in basis points). Similarly, the contribution of institutions is the sum of their allocations, from Table I, times the average of their investment management costs, from Table III. The contribution of hedge funds is the dollar cost of hedge fund and fund of fund fees, in Table IV, divided by total market cap, and the contribution of trading costs is the dollar cost, in Table V, divided by total market cap. The four components of the standardized cost and the total standardized cost are in basis points.

		Mutual Funds		Institutions		Standardized Cost				
	Market Cap	Allocation	Fees	Allocation	Fees	Mutual Funds	Institutions	Hedge Funds	Trading	Total
1980	1,103	5.2	208	46.9	34	11	16		55	82
1981	1,269	4.9	225	49.1	33	11	16		46	73
1982	1,166	5.5	196	52.1	32	11	17		55	82
1983	1,635	6.6	185	53.9	32	12	17		60	89
1984	1,639	7.3	192	55.4	31	14	17		49	80
1985	1,853	7.9	182	56.7	30	14	17		53	84
1986	2,335	9.8	180	52.8	31	18	16		54	87
1987	2,720	11.0	178	53.0	29	20	16		53	88
1988	2,470	10.4	187	50.3	30	19	15		43	78
1989	2,824	11.1	173	50.7	28	19	14		43	76
1990	2,837	11.4	165	52.9	28	19	15		38	71
1991	3,210	11.4	151	53.2	27	17	14		43	74
1992	3,771	13.3	152	53.1	28	20	15		42	77
1993	4,344	16.6	144	52.8	29	24	15		44	83
1994	4,624	19.0	143	53.4	30	27	16		40	83
1995	5,381	20.7	137	52.0	29	28	15		41	85
1996	6,881	23.3	132	48.8	25	31	12	4	38	85
1997	8,768	24.5	125	45.1	24	31	11	6	34	82
1998	10,864	25.4	119	43.6	23	30	10	4	30	75
1999	13,235	25.7	116	37.6	22	30	8	8	30	76
2000	15,675	25.6	117	37.5	24	30	9	4	32	75
2001	13,068	24.8	114	38.4	25	28	10	5	28	71
2002	11,288	25.3	112	41.3	27	28	11	6	30	75
2003	10,814	27.3	109	41.5	23	30	10	13	29	81
2004	13,183	30.0	104	41.5	22	31	9	10	25	75
2005	14,324	31.7	98	40.6	24	31	10	10	21	72
2006	15,450	34.0	95	40.0	23	32	9	13	21	75

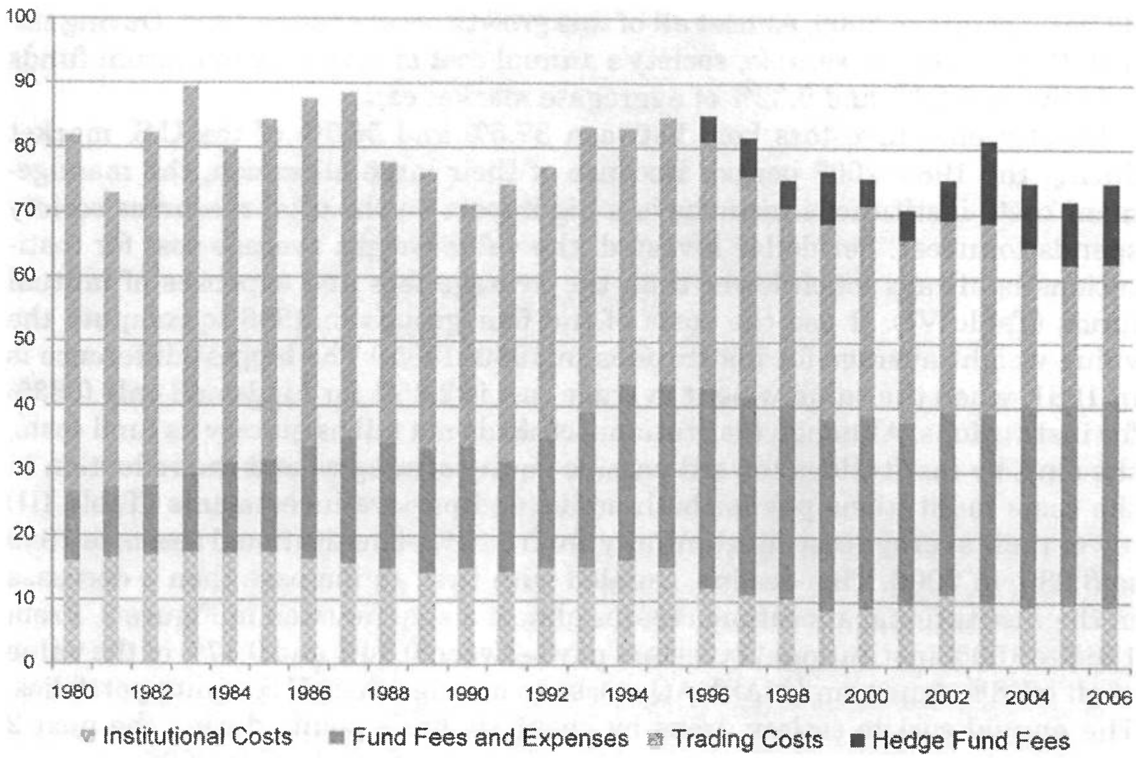


Figure 2. Fees, expenses, and trading costs relative to aggregate market cap, in basis points, 1980–2006.

of NYSE, Amex, and NASDAQ stocks during the year. The components of the standardized cost are in Figure 2.

A. The Total Cost of Investing

The average of the total standardized costs for 1980–2006, in Table VI, is 79 basis points. On average, society spends 0.79% of the aggregate value of U.S. equity to invest each year. Although the path is not smooth, the sum of the four components in Figure 2 falls gradually over time. The investment process consumes 0.82% of total market cap in 1980 and 0.75% in 2006.

There is a much larger drop in the standardized cost of trading. Investors spend 0.55% of the value of NYSE, Amex, and NASDAQ stocks to trade in 1980 and only 0.21% in 2006. Thus, during a 26-year period in which annual turnover grows from 42% to 173%, the trading revenue of brokers and dealers declines from about two-thirds of society's total cost of investing to less than one-third.

Much of the decline in trading costs is offset by an increase in the cost of mutual funds. Driven mostly by falling open-end loads, the value-weight average cost per dollar invested in U.S. equity funds, in Table VI, drops fairly steadily from a stiff 2.08% in 1980 to 0.95% in 2006. But the allocation to mutual funds increases by more, from only 5.2% of U.S. equity in 1980 to 34.0% in 2006. The net result is a tripling of the standardized cost, from 11 basis points in 1980 to

32 basis points in 2006. Almost all of this growth occurs before 1995. During the last 12 years of the sample, society's annual cost of investing in mutual funds is between 0.28% and 0.32% of aggregate market cap.

Institutional investors hold between 37.5% and 56.7% of the U.S. market during the 1980–2006 period. Because of their large allocation, the management costs institutions incur have a big impact on the total resources society spends to invest. Per dollar invested, the value-weight average cost for institutions is always much lower than the average fees and expenses of mutual funds (Table VI). (I use the costs of the four groups in 1986 to compute the value-weight average for institutions in 1980–1985.) The biggest difference is in 1981, when the value-weight average cost is 2.25% for funds and only 0.33% for institutions. Although institutional costs do not fall as quickly as fund costs, the shift by institutions toward passive equity strategies and the reduction in the costs institutions pay for both active and passive investments (Table III) lower their average cost substantially, from 0.34% of institutional assets in 1980 to 0.23% in 2006. This decline, coupled with first an increase then a decrease in the institutional allocation, creates almost a step function in Figure 2. From 1980 to 1995, institutional investors pay between 0.14% and 0.17% of the value of all NYSE, Amex, and NASDAQ stocks to manage their U.S. equity portfolios. The annual cost to society drops by about six basis points during the next 2 years, and remains between 8 and 11 basis points from 1997 to 2006.

If we ignored hedge funds, Figure 2 would say that society's cost of investing in U.S. equity falls a lot over time, from 0.82% of aggregate market cap in 1980 to 0.62% in 2006. Although the big shift from direct holdings to mutual funds pushes up the cost of investing, this effect is overwhelmed by the reduction in trading costs and, to a lesser extent, the decline in institutional management costs. The net effect would be a 24% reduction in the cost of investing per dollar of stock market wealth.

But we cannot ignore hedge funds. Measured in dollars, hedge fund and fund of fund fees on U.S. equity-related assets jump from 2.8 billion in 1996 to 19.4 billion in 2006 and 25.0 billion in 2007 (Table IV). The standardized cost is equally impressive. Fees on U.S. equity-related hedge fund assets, in Table VI, grow from 0.04% of the value of the market in 1996 to 0.13% in 2006. Given the relatively small size of the industry, these seem like big numbers. The fees hedge fund and fund of fund clients pay to invest 458.6 billion dollars in 2006, for example, are 36% higher than all the costs institutions pay to invest 6.18 trillion.

Hedge fund fees absorb about two-thirds of the reduction in the other costs of investing, but they do not claim them all. Though the process is not smooth, the total cost of investing—including hedge fund fees—falls from 0.82% of aggregate market cap in 1980 to 0.75% in 2006.

B. The Cost of Investing if Everyone Is Passive

Passive investors incur some costs. Thus, the incremental cost of active investing is the difference between society's total cost, in Table VI and Figure 2, and the resources that would be consumed if all investors followed a passive strategy. I make several assumptions to estimate society's cost of investing passively. First, investors in mutual funds switch to a passive mutual fund whose cost matches the highest expense ratio among the share classes of Vanguard's Total Stock Market Index, an open-end fund that holds NYSE, Amex, and NASDAQ stocks.[4] (Vanguard started the fund in 1992, so I use the expense ratio for that year as the cost before 1992.) Second, institutions also move their U.S. equity investments to a passive market portfolio. For most institutions, the cost of this portfolio is the cost of the passive DB investments monitored by CEM Benchmarking. Defined contribution plans pay a bit more. As in the estimates in Table III, their cost is the passive DB cost plus the average difference between the costs of passive DC and DB plans. (I use the 1991 estimates for 1980–1990.) Third, I continue to assume that there are no fees and expenses associated with direct holdings and ESOPs. Fourth, in the passive scenario hedge fund investments are reallocated proportionately among direct holdings, mutual funds, and institutions.

Finally, I assume that if all investors follow a passive strategy, total turnover is 10% a year. This assumption has a big impact on my results and, because the cost of trading declines over time, the impact is bigger early in the period. Lowering the assumed turnover to 5%, for example, cuts my estimate of the cost of passive investing by 6.7 basis points in 1980 and only 0.5 basis points in 2006. Passive investors trade for two reasons, to accommodate cash flows and to maintain target risk-return tradeoffs. When thinking about the appropriate turnover for the passive scenario, it is important to remember that a large slice of the market would be held by passive institutions with only modest inflows and outflows. Moreover, most flows from mutual fund clients would cross at the fund level, without any need for trading. Although a lower passive turnover may be appropriate, the 10% assumption is conservative because it pushes up the estimated cost of passive investing and lowers my estimate of the resources investors spend to beat the market.

The results of these calculations are in Table VII. The components of society's cost of investing in the passive scenario are muted versions of the actual costs in Table VI. Because of the shift from direct holdings to mutual funds (Table I), the standardized cost of mutual funds increases from 1.1 basis points in 1980 to 6.5 basis points in 2006. The 60% reduction in the institutional cost of passive investing (Table III) and the modest reduction in the allocation to institutions over time (Table VI) combine to lower institutional costs from 3.6 basis points to only 1.2 basis points. Though not surprising, the drop in trading costs is most dramatic. The standardized cost in 1980, 13.3 basis points, is 11 times the cost of 1.2 basis points in 2006. The net result is a 50% reduction in the standardized cost of passive investing, from 0.180% of the value of all NYSE, Amex, and NASDAQ stocks in 1980 to 0.089% in 2006.

[4] Vanguard holds large stocks in proportion to their market caps, but it samples small stocks, overweighting some and holding no shares of others. Sampling reduces the fund's custodial costs and expense ratio. Its impact on trading costs is ambiguous.

Table VII
Standardized Cost of Passive Investing, in Basis Points, and Incremental Cost of Active Investing, in Basis Points and Billions of Dollars, 1980–2006

The standardized cost of passive would be the cost of investing if all U.S. equity were held passively and is measured relative to the market cap of all NYSE, Amex, and NASDAQ stocks. Actual – Passive, the incremental cost of active in basis points, is the standardized cost of investing (Table VI) minus the passive cost. The average of the annual differences is reported for 1980–2006. Price discovery, the incremental cost of active in billions of dollars, is Actual – Passive times the aggregate market cap.

	Standardized Cost of Passive				Incremental Cost of Active	
	Mutual Funds	Institutions	Trading	Total Cost	Actual – Passive	Price Discovery
1980	1.1	3.6	13.3	18.0	64	7.0
1981	1.0	3.7	11.7	16.5	56	7.2
1982	1.1	3.9	10.8	15.8	67	7.8
1983	1.4	4.0	10.6	15.9	74	12.0
1984	1.5	3.9	8.7	14.2	66	10.8
1985	1.7	4.0	8.1	13.8	70	13.1
1986	2.1	3.7	7.2	13.0	74	17.4
1987	2.3	3.8	6.3	12.4	76	20.7
1988	2.2	3.6	6.6	12.4	65	16.1
1989	2.3	3.6	6.4	12.3	64	18.0
1990	2.4	3.8	6.4	12.6	59	16.6
1991	2.4	3.8	7.0	13.2	61	19.6
1992	2.7	2.9	6.9	12.4	65	24.4
1993	3.3	3.2	6.4	13.0	70	30.5
1994	4.7	3.6	5.7	14.1	69	31.9
1995	4.1	2.9	5.1	12.2	73	39.1
1996	4.7	2.5	4.5	11.6	73	50.5
1997	4.9	2.2	3.7	10.7	71	62.1
1998	5.1	2.0	3.2	10.2	65	70.5
1999	5.1	1.4	2.7	9.3	67	88.8
2000	5.1	1.6	2.3	9.1	66	103.9
2001	5.0	1.7	2.3	9.0	62	80.4
2002	5.1	1.7	2.1	8.9	66	74.8
2003	5.5	1.2	2.2	8.9	72	78.3
2004	5.7	1.2	1.8	8.6	66	87.3
2005	6.0	1.2	1.4	8.6	63	90.7
2006	6.5	1.2	1.2	8.9	66	101.8
1980–2006					67	

C. The Cost of Active Investing

We are now ready to answer the central question. The average difference between the actual standardized cost of investing and the passive cost for the 1980–2006 period, in Table VII, is 67 basis points. On average, active investors spend 0.67% of the total market cap each year on what, in aggregate, is a futile search for superior returns. If we assume that society will continue to spend the current real dollar cost of active investing forever and that the expected real return on the U.S. stock market is a constant 6.7%, the capitalized cost is 10% of the current value of the market. This estimate is conservative. First, the estimates in Fama and French (2002) and Graham and Harvey (2005) suggest that the long-term equity risk premium is far below 6.7%. If so, the expected real return on the market is almost certainly below 6.7%. Second, the data imply that the annual dollar cost of active investing will grow with the aggregate market cap. Positive expected growth and a lower discount rate both push the capitalized cost above 10%. In short, if the social benefit of active investing is price discovery, the annual cost is 0.67% of the aggregate value of the market and the capitalized cost is at least 10% of the value.

Figure 3 plots the difference between the actual and passive costs of investing. Standardized by aggregate market cap, the cost of active investing is remarkably stable. All but 3 of the 27 estimates for 1980–2006—including all of the estimates after 1990—are between 61 and 74 basis points. There is also little evidence of a time trend in the incremental cost. The average difference between the actual and passive costs in Table VII is 66 basis points for the first half of the period and 68 basis points for the second half. Of course, the lack of a time trend is driven in part by the assumption of 10% turnover in the passive scenario. If passive turnover is 5%, the standardized cost of trying to beat the U.S. stock market falls by three basis points from 1980 to 2006, and if passive turnover is 15%, the standardized cost rises by eight basis points over the period.

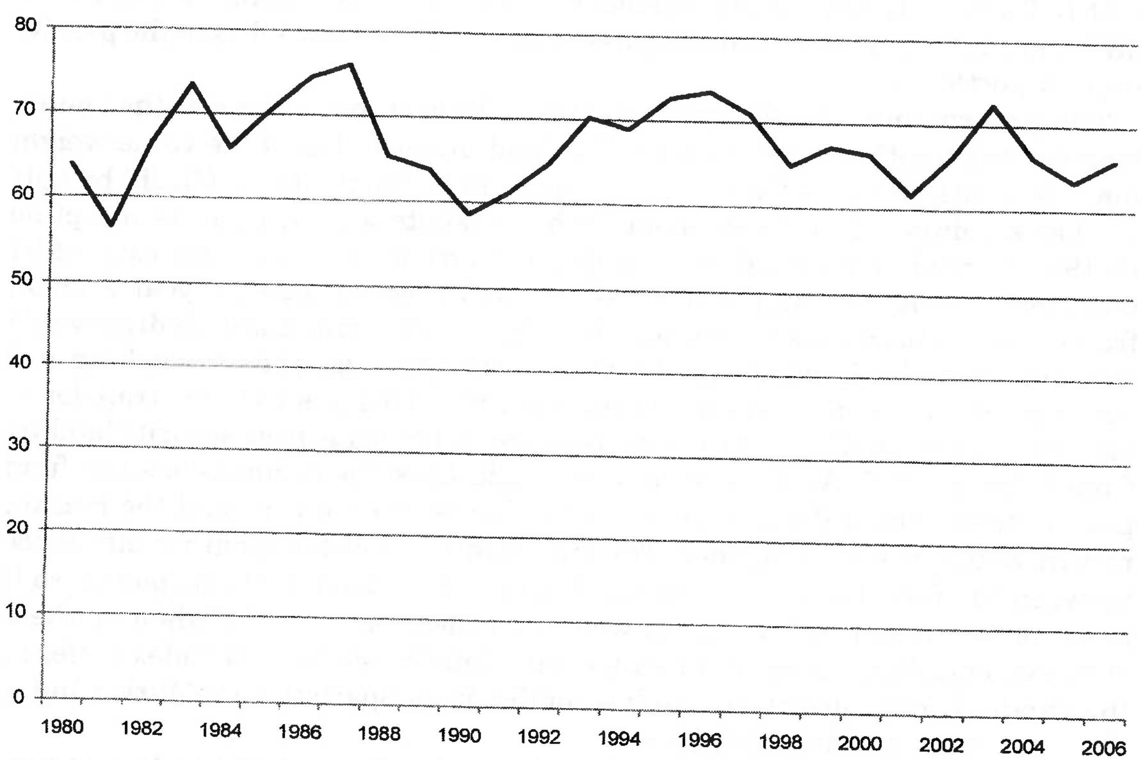


Figure 3. The difference between the actual and passive costs of investing, in basis points, 1980–2006.

Table VII also reports the dollar cost of active investing. This is the aggregate market cap (from Table VI) times the difference between the standardized actual and passive costs of investing. With 10% passive turnover, the incremental cost per dollar invested is relatively constant over time, so the total dollar cost grows with the market. The cost of active investing is 7.0 billion dollars in 1980, 30.5 billion in 1993, and 101.8 billion in 2006. Thus, in 2006 investors searching for superior returns in the U.S. stock market consume more than 330 dollars in resources for every man, woman, and child in the United States.

Finally, the results in Table VII allow me to address a closely related question. How would a small but representative investor's return change if he switched from the value-weight combination of all investors' strategies to a passive market portfolio? Because the combination of all investors' portfolios is the market portfolio, the representative investor's initial return is the gross return on the market minus the value-weight average of all investors' costs. Any trading gains, losses, and other transfers between investors happen within his portfolio and have no effect on his return. This is not the case if the representative investor switches to a passive market portfolio. Trading gains and security lending fees paid by active investors to borrow shares from the passive portfolio, for example, push up his return and trading losses lower it. Thus, to use the cost of investing in the passive scenario to measure the return when the representative investor switches, I have to assume there is no net transfer between the passive market portfolio and other investors. With this assumption, the return on the passive market portfolio is the gross return minus the passive cost in Table VII, and the representative investor increases his return by the difference between the actual and passive costs when he switches to the passive market portfolio.

The performance of Vanguard's Total Stock Market Index suggests the no-net-transfer assumption is reasonable. The fund underperforms the value-weight market return from the Center for Research in Security Prices (CRSP) by only 2.1 basis points a month, or about 25 basis points a year, from its inception in 1992 to September 2006. If we add the fund's average expense ratio of 21 basis points to its return, the shortfall drops to 4 basis points per year. A small fraction of the fund's assets is typically in cash. Reversing this cash drag would add another 7 basis points to the fund's average gross return, pushing it 3 basis points *above* the average market return. This positive net transfer is almost exactly equal to the fund's average annual revenue from security lending from 1998 to 2007. And if I were able to add back the commissions the fund pays to trade, the difference between the fund's gross return and the market return would rise even further. The standard error of the monthly difference between the fund return and the market return is about 1.5 basis points, so it is important to not put too much weight on these results. Nonetheless, there is no evidence that investors in Vanguard's Total Stock Market Index suffer at the hands of active investors. Analysis of Fidelity's Spartan Total Market Index Fund produces a similar conclusion.

The evidence from the Vanguard and Fidelity funds suggests the no-net-transfer assumption is conservative. Thus, it seems safe to use the cost of investing in the passive scenario to estimate the returns one could earn on a passive market portfolio. If so, a representative investor who switches to a passive market portfolio would increase his average annual return by 67 basis points from 1980 to 2006.

VII. Summary and Conclusions

I compare the resources society spends to invest in the U.S. stock market with what would be spent if everyone followed a passive strategy. My estimate of the actual cost of investing—the fees and expenses paid for mutual funds, the investment management costs paid by institutions, the fees paid to hedge funds and funds of funds, and the transaction costs paid by all traders—is 0.82% of the value of all NYSE, Amex, and NASDAQ stocks in 1980 and 0.75% in 2006. In the passive scenario, investors pay passive fees, annual turnover is 10%, and there are no hedge funds. As a result, the cost of investing is only 0.18% of the aggregate market cap in 1980 and 0.09% in 2006.

The difference between the actual and passive estimates measures the cost of active investing. The average difference for 1980–2006 is 0.67%. Thus, from society's perspective, the average annual cost of price discovery is 0.67% of the total value of domestic equity and the capitalized cost is at least 10% of the current market value. From a typical investor's perspective, the message is more challenging. If there is no net transfer between a passive market portfolio and other investors, the average annual return on the passive portfolio is 67 basis points higher than the value-weight average of all investors' returns. Thus, if a representative investor switched to a passive market portfolio, he would increase his average annual return by 67 basis points over the 1980–2006 period.

Hedge fund fees in 2007 are a stark illustration of the negative sum nature of active trading. The value-weight average fee on U.S. equity-related hedge fund assets in 2007 is 4.63% and the average fund of fund fee is 1.85%. Since fund of fund investors must pay both their own funds' fees and the fees of the underlying hedge funds, the typical fund of fund investor does not break even in 2007 unless U.S. equity-related hedge funds generate average abnormal returns of 6.48%. There are 458.6 billion dollars invested in hedge funds at the beginning of 2007, so even if we ignore the other costs they incur, hedge funds must take 29.7 billion dollars in abnormal profits from other U.S. equity investors for their fund of fund clients to break even. The total capitalization of the U.S. market is 16.53 trillion dollars at the beginning of 2007, so a 29.7 billion dollar transfer would reduce the value-weight average return of all nonhedge fund investors by about 18 basis points. Of course, if passive investors do not participate in the transfer the burden for active investors is even higher. They must contribute about 22 basis points of their U.S. equity holdings in 2007 for fund of fund investors to break even. And these losses would be on top of the active investors' own fees, expenses, and trading costs.

Whether fund of fund investors break even or not, a passive market portfolio produces a higher return than the aggregate of all active portfolios. Why do active investors continue to play a negative sum game? Perhaps the dominant reason is a general misperception about investment opportunities. Many are unaware that the average active investor would increase his return if he switched to a passive strategy. Financial firms certainly contribute to this confusion. Although a few occasionally promote index funds as a better alternative, the general message from Wall Street is that active investing is easy and profitable. This message is reinforced by the financial press, which offers a steady flow of stories about undervalued stocks and successful fund managers.

Overconfidence is probably the other major reason investors are willing to incur the extra fees, expenses, and transaction costs of active strategies. There is evidence that overconfidence leads to active trading. (See, for example, Odean (1998), Barber and Odean (2001), and Statman, Thorley, and Vorkink (2006).) Investors who are overconfident about their ability to produce superior returns are unlikely to be discouraged by the knowledge that the average active trader must lose.

Statman (2004) offers another behavioral explanation for active investing. He suggests that, in addition to expected return and risk, investors are concerned with what he calls the expressive characteristics of their portfolios. Thus, some investors may accept a lower expected return in exchange for the bragging rights that come with a fund that has performed well. Others may give up the low cost and diversification of a passive mutual fund for the prestige of their own separate account.

Finally, some investors trade actively because they really are able to produce superior returns. The existence of superior investors, however, does not explain the behavior of the average investor. Active investing is still a negative sum game. Every dollar a superior investor earns must increase the aggregate losses of all other active investors.

Appendix

A. Allocation of Equity

The main source for the allocation of U.S. publicly traded common equity in Table I is the December 6, 2007 release of the Federal Reserve Board's Flow of Funds Accounts. Table L.213 of the Flow of Funds Accounts reports the value of corporate equity held by various groups of investors, such as households and nonprofits, mutual funds, and insurance companies. The Fed uses the household and nonprofit sector as a residual. Its allocation is the aggregate value of corporate equity minus the combined values of the other sectors. Thus, the household and nonprofit sector includes not only the publicly traded common equity held by households and nonprofits, but also preferred stock and closely held corporations. Many of the calculations in this section are to separate these pieces.

I start by eliminating preferred stock. In personal correspondence, Standard and Poor's generously provided estimates of the total value of preferred stock from their internal stock and bond database for most of the years from 1980 to

2007.[5] I use exponential interpolation to fill in the missing years, 1981–1985 and 1987. To expedite the discussion, below I refer to what remains in the household and nonprofit sector after subtracting preferred stock as simply the value of the household and nonprofit sector.

The Federal Reserve reports separate estimates of the holdings of nonprofits for 1988–2000 (Flow of Funds Accounts table L.100a) and I use them to calculate the allocations for those years in Table I. I estimate the holdings of nonprofits in each year before 1988 as the value of the household and nonprofit sector for the year times the 1988 ratio of nonprofit holdings to the value of the household and nonprofit sector, and I use the ratio for 2000 to estimate nonprofit holdings for 2001–2007.

The direct holdings of households in Table I build on estimates in Kennickell (2003, 2006). He uses information in the Fed's triennial Survey of Consumer Finances (SCF) to measure the amount of publicly traded equity households own directly in 1989, 1992, 1995, 1998, 2001, and 2004. Kennickell's estimates are adjusted for inflation. After converting them back to nominal dollars, I adjust his estimates by the annual value-weight average return on U.S. stocks, from CRSP, to infer the value of direct holdings for the missing years between 1989 and 2004. My estimate for 2002, for example, is the nominal value of Kennickell's estimate for 2001 times one plus the market return for 2002 and my estimate for 2003 is the 2004 value divided by one plus the market return for 2004. I estimate the value of direct holdings for each year before 1989 and after 2004 as the value of the household and nonprofit sector for the year times either the 1989 or 2004 ratio of direct holdings to total household and nonprofit holdings. I assume Kennickell's measure of households' direct holdings of publicly traded equity includes the value of ETFs. Since the Flow of Funds Accounts has a separate allocation for ETFs, I reduce my estimate of direct holdings by the Fed's estimate of the value of ETFs.

Although investment costs differ across DB plans, DC plans, and ESOPs, the Federal Reserve combines their allocations in table L.213 of the Flow of Funds Accounts. The Fed does report the value of U.S. equity held by DB plans and DC plans in 1985–2006 in tables L.118b and L.113c. To estimate the U.S. equity held by DB plans in 2007, I assume they do not change the ratio of their holdings of U.S. equity relative to all private pension plan assets from 2006 to 2007. The Department of Labor's website reports the total assets in DB plans and in DC plans (including ESOPs) for 1975–2005.[6] I assume the DB plans' share of the U.S. equity held by private pensions in 1980–1984 is proportional to their share of the total assets in private pensions. Thus, to estimate the U.S. equity held by DB plans in 1980–1984, I multiply the total allocation to private pensions in table L.213 by the ratio of the total assets in DB plans divided by the total assets in DB and DC plans.

Both the Federal Reserve and the Department of Labor combine ESOPs with other DC plans in their estimates. The annual survey results for 1980–2006

[5] I thank Shrikant Dash for this information.

[6] The information is at www.dol.gov/ebsa/pdf/privatepensionplanbulletinhistoricaltables.pdf.

from Greenwich Associates include the value-weight average fraction of DC plan assets allocated to a company's own stock. I use this fraction to separate ESOPs from other DC plans. Thus, I start by estimating the total assets in DC plans for 1980–2007 using the approach I describe for DB plans above. I then use the data from Greenwich Associates to split ESOPs from other DC plans. (The 2007 split uses the Greenwich estimate for 2006.) To be clear, the allocation to DC plans in Table I does not include ESOPs.

The Federal Reserve's allocations in table L.213 include the foreign equity held by U.S. investors. Thus, my next step is to remove these securities by assuming they are held proportionately by all U.S. investors except ESOPs.

The Flow of Funds Accounts do not include a separate allocation to hedge funds. As I describe in Section D below, I use estimates of the total assets invested in hedge funds, from HFR, to compute hedge fund and fund of fund fees. To avoid double counting, however, I have to reduce the U.S. equity allocations of other groups of investors by the net holdings of hedge funds. Because hedge funds invest in a variety of assets, hold short and long positions, and use leverage, their net holdings of U.S. equity differ from the total assets invested. My indirect measure of net holdings multiplies HFR's estimate of total hedge fund assets by the slope coefficient from a regression of hedge fund returns on U.S. market returns. For example, a 100 million dollar portfolio with long and short U.S. equity positions of 250 and 200 million dollars has net holdings of 50 million and a slope on the U.S. market of about 0.5.

I estimate the aggregate slope for all hedge funds by regressing the monthly return (in excess of the U.S. Treasury bill rate) on the CSFB/Tremont Hedge Fund Index, a broad value-weight index of hedge funds, against the excess returns on the CRSP value-weight index of NYSE, Amex, and NASDAQ stocks. To control for correlations with other markets, I also include excess returns on MSCI's Emerging Markets and World Ex-U.S. (developed markets) Indices in the regression. The estimated slopes (and standard errors) for January 1994, the start of the CSFB/Tremont index, to September 2007 are given by:

$$R_{HF,t} = \underset{(0.11)}{0.44} + \underset{(0.06)}{0.19}\, R_{US,t} - \underset{(0.06)}{0.04}\, R_{Dev,t} + \underset{(0.03)}{0.11}\, R_{Emrg,t} + et. \qquad \text{(A1)}$$

In this regression, $R_{HF,t}$ is the hedge fund return in month t and $R_{US,t}$, $R_{Dev,t}$, and $R_{Emrg,t}$ are the returns on the U.S., developed markets, and emerging markets indices. The adjusted regression R^2 is 0.34.

Asness, Krail, and Liew (2001) find that the returns on some components of the CFSB/Tremont index are correlated with lagged market returns during their 1994–2000 sample. When I add lagged market returns to regression (A1), however, the lagged slopes are indistinguishable from zero and the contemporaneous slopes are essentially unaffected. Thus, I use regression (A1) to measure the sensitivity of hedge fund returns to the U.S. market. In short, the allocation to hedge funds in Table I is 19% of the total hedge fund assets reported by Hedge Fund Research.

The Federal Reserve uses cross-border flows to estimate the U.S. equity held by foreign investors. As a result, the U.S. equity held by foreign-domiciled hedge funds is included with the holdings of other foreign investors. The U.S. equity held by hedge funds domiciled in the U.S. is in the Fed's residual category, households and nonprofits. I use individual fund data from HFR to measure the fraction of hedge fund assets domiciled in the U.S. I remove that fraction of the net holdings of hedge funds from my estimate of direct holdings and I subtract the rest from the holdings of foreign investors.

Finally, I allocate the U.S. equity holdings of foreign investors proportionately among direct holdings, mutual funds, closed-end funds, ETFs, DB plans, and banks, insurance companies, and broker/dealers. This allocation excludes nonprofits, DCs plans, ESOPs, public plans, and state and local governments.

B. Mutual Funds

The data used to compute the value-weight average mutual fund fees in Table II are from the September 2006 version of the Survivor Free Mutual Fund Database from the Center for Research in Security Prices at the University of Chicago. The major challenge is identifying U.S. equity funds. I use S&P objective codes, policy codes, area codes, Weisenberger fund types, and fund names to classify funds. A fund's name and style codes can change over time. I exclude a fund during any period in which I cannot infer that its assets are both domestic and primarily equity.

The S&P objective code is not available until 1993, the area code begins in July 2003, and the policy code is not available after 1990. The Weisenberger code is imprecise. Thus, although it begins earlier, I use it only during 1991, 1992, and 1993, when no other style codes are available.

The process I use to infer the nature of a fund's assets from its name is based on a mapping from 977 character strings to 78 investment styles. Municipal bond funds in the CRSP database, for example, typically have "Municipal," "Muni," or "Mu Tr" in their names. Including different capitalizations, I identify 22 strings associated with small cap value and 2 for small/mid value. This mapping has many exceptions. "High yield," for example, usually signals a bond fund, but not if it is followed by "stock." Similarly, none of the funds with "Barclays Global" in the name are actually global. The algorithm to interpret fund names has more than 250 overrides for specific cases like these.

I try to determine whether a fund is definitely equity and definitely domestic during each month it is in the database. Sometimes the style codes and fund name contradict each other. The S&P objective code appears to be the most reliable so a definite classification based on this code trumps almost all other information. For example, if the S&P code says a fund is definitely not equity in 2001, I exclude the fund from my calculations for that year. I override the S&P code only if the fund's current name implies its assets are definitely not equity or definitely not domestic. If the S&P code is not available or does not reveal the investment region, I turn to the area code. The Weisenberger code is next, followed by fund name. Finally, I use the policy code for any month in which the fund's region or asset class remains uncertain. In short, I look at a fund's S&P objective code, area code, Weisenberger code, name, and policy code sequentially each month. I include the fund in the sample only if the codes and fund name say that the fund is definitely domestic equity before they say it is definitely not equity or definitely not domestic.

The average mutual fund expense ratios in Table II weight funds by their assets under management at the beginning of the year. Replacing missing expense ratios with the equal-weight average expense ratio of funds of similar size has a negligible effect on the results.

The average annuitized loads for mutual funds in Table II are from the Investment Company Institute and are described in Rea and Reid (1998).[7] The averages I use weight funds by their sales. Switching to asset-weight averages increases the average annuitized load—and my estimate of society's cost of trying to beat the market—by an average of one basis point a year.

C. Hedge Fund Fees

Hedge fund and fund of fund managers often charge two fees, a management fee that is a fixed percent of current assets and a performance fee that depends on the fund's profits. Funds usually pay the management fee more frequently, but the performance fee is almost always paid only once a year, typically at the end of December. The performance fee may depend on a high water mark or a hurdle rate, which may be a constant, such as 10%, or the return on a financial instrument, such as 1-month Treasury bills. To understand how high water marks and hurdle rates affect performance fees, define a fund's adjusted gross return for a year as its gross return minus its management fee. If there is a hurdle rate and no high water mark, the annual performance fee is a function of the maximum of zero and the difference between the current adjusted gross return and the hurdle rate; the fee depends on only this year's return. A high water mark puts memory in the process. With a high water mark, the annual performance fee for a new investor is proportional to the maximum of zero and the difference between the cumulative adjusted gross return since he invested and the cumulative hurdle rate. The annual performance fee for an investor who has paid at least one fee is the maximum of zero and the difference between the cumulative adjusted return since his last performance fee and the cumulative hurdle rate.

Management and performance fees accrue until they are paid. Most funds report their monthly net return, which is the gross return minus the change in the accrued fees for an investor who was in the fund the last time a performance fee was paid. Although the realized performance fee is one-sided—the manager does not contribute money if the fund does poorly—accrued performance fees can be recovered. Thus, if a fund starts the month with a positive accrued performance fee and then performs poorly, the fund's net return is increased by a reduction in the accrued fee.

[7] Sean Collins of the ICI kindly provided the asset-weight and sales-weight averages of the annuitized load for 1980–2006.

Because high water marks make performance fees a function of past returns, they complicate calculations to convert net returns into gross returns and performance fees. Fortunately, the link with past returns is broken when a performance fee is paid. The relations among net returns, gross returns, quoted fees, and actual fees for a fund with a high water mark depend on only the fund's returns since its most recent positive performance fee. In the equations below, I denote the date of that fee as time zero, $t = 0$, and I assume management fees and performance fees are paid yearly, so each period is 1 year.

Define M as the fund's annual management fee (e.g., 2%) and P as the quoted performance fee (e.g., 20%). Also define $H(t)$ as one plus the hurdle rate for year t, $G(t)$ as one plus the gross return, and $G'(t)$ as the adjusted gross return, $G'(t) = G(t) - M$. Finally, define $N(t)$ as the compounded value of the adjusted gross return, $N(0) = 1$ and $N(t) = N(t-1) * G'(t)$.

An investor does not pay a performance fee in year t unless the value of his investment before subtracting the fee is above the high water mark. Consider an investor with one dollar in the fund at time $t = 0$. His high water mark is the compounded hurdle rate, $HWM(0) = 1$ and $HWM(t) = HWM(t-1) * H(t)$, and his net return in year t is the adjusted gross return, $G(t) - M$, minus the performance fee. If the next positive performance fee is in year T, his investment is worth the compounded value of the adjusted gross return, $N(t)$, at the end of each year before T and it is worth $N(T)$ before subtracting the performance fee in T. Thus, for each dollar invested at time 0, the performance fee in year t is $P * \text{Max}[0, N(t) - HWM(t)]$. Since his investment is worth $N(t-1)$ at the beginning of year t, the net return for year t is

$$\begin{aligned} R(t) &= G(t) - M - P * \text{Max}[0, N(t) - HWM(t)]/N(t-1) \\ &= G'(t) - P * \text{Max}[0, G'(t) - hwm(t)], \end{aligned} \tag{A2}$$

where $hwm(t) = HWM(t)/N(t-1)$ is the high water mark at the end of year t relative to the investment at the beginning of t.

I use a sequential process to convert the net returns firms typically report into gross returns and realized performance fees, $P * \text{Max}[0, G'(t) - hwm(t)]$. I assume each fund has just paid a performance fee when it is added to the database. I then compare the net return and the relative high water mark for each successive year t. If the net return is less than the relative high water mark, the fund does not pay a performance fee and the gross return is the net return plus the management fee. And if the net return is greater than the relative high water mark, the fund did pay a performance fee, the gross return is the net return plus both fees, and I restart the process.[8]

I use the December 2007 version of Hedge Fund Research's live and graveyard databases to measure hedge fund and fund of fund fees. The live database

[8] Suppose the adjusted gross return is less than the relative high water mark, $G'(t) < hwm(t)$. Then equation (A2) implies $R(t) = G'(t)$ and the net return is also less than the relative high water mark. Similarly, if the adjusted gross return is greater than the relative high water mark, $G'(t) \geq hwm(t)$, we can rewrite (A2) as $R(t) - hwm(t) = (1 - P)\ [G'(t) - hwm(t)]$ and, since $P < 1$, the net return is greater than the relative high water mark.

contains funds that are currently active and willing to have their performance reported to HFR's clients. The graveyard database contains historical data on dead funds and on active funds that withdraw from the live database. HFR maintains a private database of active funds that are not in the live database. The information in this database, which is from a variety of sources including fund of fund managers, other fund investors, and the funds themselves, is used to estimate aggregate hedge fund assets and the performance of hedge fund indices.

The HFR data have several virtues. First, the graveyard database minimizes survival bias. Second, HFR records the date each fund is added to the databases, so it is easy to avoid backfill bias. Third, HFR reports details of each fund's fee, including whether there is a high water mark or a hurdle rate and, if there is a hurdle rate, how it is set.

Despite these virtues, the HFR data are not perfect. HFR reports only the most recent fee for each fund. Funds rarely change their quoted fees, however, so this is not a big problem. More important, the fees in the databases are quoted prices, not the contractual fees investors actually pay. Since deals with individual clients are private, this problem afflicts every study of hedge funds, but it may not be severe. Total assets invested in hedge funds grow rapidly during the last seven years of the sample, from less than 500 billion dollars at the beginning of 2001 to 1.81 trillion dollars in 2007. Some industry experts suggest that, on a value-weight basis, actual fees are not far from quoted fees, particularly during the period of explosive growth when the demand for access to funds forces many if not most investors to pay list price.

The HFR's public databases are also not comprehensive. A fund is included only if the manager chooses to provide the necessary information. If the manager stops reporting, HFR searches for a final return and moves the fund to its graveyard database. Inclusion in the (live) database is perceived to be helpful to managers who are trying to raise assets. Thus, the database is probably biased toward younger and smaller funds. It is not clear how the tilt away from more established funds affects average returns, but it probably pushes the sample toward funds with higher return variances and realized performance fees.

The live and graveyard databases report monthly performance and assets under management. They also report: (i) the current management and performance fees for live funds or the last fees for graveyard funds; (ii) whether the fund has a high water mark; (iii) whether the fund has a hurdle rate and, if so, how the hurdle rate is determined; (iv) whether the reported returns are net of all fees, net of only the management fee, or gross of fees; (v) whether the fund is domiciled outside the U.S.; and (vi) the currency in which the returns and assets under management are denominated.

Since my goal is to estimate the resources spent trying to produce superior returns in the U.S. stock market, I want to measure only the hedge fund and fund of fund fees paid for U.S. equity-related investments. I use categories assigned by HFR to sort funds into three groups. I assume funds in the first group use 100% of their assets for equity trading strategies, those in the second use 50%, and those in the third do not use any of their assets for equity trading.

100% Equity	50% Equity	No Equity
Convertible Arbitrage	Distressed Securities	Emerging Markets
Equity Hedge	Event Driven	Fixed Income
Equity Market Neutral	Macro	Managed Futures
Equity Nonhedge	Market Timing	Foreign Exchange
Merger Arbitrage	Regulation D	
Sector Funds	Relative Value Arbitrage	
Short Selling		

I assume Regulation D funds invest only in the U.S., but other funds invest around the world. Since HFR has a separate category for emerging markets funds, I use the weight of the U.S. in the portfolio of all developed market equities, from S&P/Citigroup, to estimate the fraction of equity-related hedge fund assets invested in the U.S.

The results summarize the fees for 3,714 hedge funds in the 50% and 100% equity categories and the graveyard database has 2,666. The databases have 2,452 and 783 funds of funds. I use exchange rates from Reuters (provided by Dimensional Fund Advisors) to convert foreign currencies to dollars. I drop one hedge fund and three funds of funds denominated in European currency units (ECU), and one hedge fund denominated in Czech krona. HFR provides information only for the most recent currency. Thus, I do not know the initial currency of six hedge funds and two funds of funds that convert to the euro when this currency is introduced in 1998.

I also drop a fund if HFR does not report either its management or performance fee. This requirement rules out 102 hedge funds and 35 funds of funds from the graveyard database and 23 hedge funds and 60 funds of funds from the live database. I drop 22 funds from the graveyard database and 38 funds from the live database because both the reported management fee and the reported performance fee are zero. Five funds in the live database and 14 funds in the graveyard database are missing at least one monthly return. I replace the missing data with zero when calculating the results in Table IV, but dropping the 19 funds completely has a negligible effect on my estimates. There are many more funds with missing assets. I do not include a fund until the first month assets are available after the fund is added to the HFR database and I assume assets grow at the fund's reported return when they are missing.

HFR uses all three of its databases to measure total assets in each category and the hedge fund assets in Table IV are based on these estimates. To estimate the U.S. equity-related fund of fund assets in Table IV, I multiply total fund of fund assets by the ratio of U.S. equity-related hedge fund assets relative to all hedge fund assets. The fees in Table IV are averages of the value-weight average for each category. Thus, I use individual fund data to compute the value-weight average for each category, then I weight each average by the category's total beginning-of-year assets times the fraction of its assets in U.S. equity-related strategies.

D. Trading Costs

Registered securities firms must file FOCUS reports with the SEC each year. These reports contain detailed financial statements, including information about the revenue firms earn by trading. I use aggregate values of these data, from the SEC, to estimate the exchange commissions, over-the-counter (OTC) commissions, and trading gains in Table V. The process I use is almost identical to the process in Stoll (1993).

The relevant FOCUS data for 1980–2006 are in Table A1. There are two versions of the reports. Firms that clear trades or carry customer accounts use Part II and those that do neither use the simpler Part IIA. The commissions and market-making gains in Table IV combine the revenues for Parts II and IIA firms.

I make three adjustments to the data in Table A1. First, the exchange and OTC commissions for equity trades include the commissions, clearing fees, and floor brokerage fees that one securities firm pays to another. These transactions are transfers, rather than an additional cost of trading, so they should be eliminated from reported commissions. The FOCUS reports show the total value of transfers between firms, but there is not a separate line for just U.S. equity transactions. Thus, to eliminate the transfers, I follow Stoll (1993) and assume the transfers for each group of trades are proportional to the commissions for those trades. For example, I reduce exchange commissions by the total value of the transfers times the ratio of exchange commissions to total commissions. Second, the FOCUS reports pool many equity commissions with commissions from other sources. The "Other" line includes all commissions except (i) those for listed options and listed equity (Part II firms), or (ii) those for listed options and listed equity traded on an exchange (Part IIA firms). I use Stoll's (1993) estimate that 90% of these "Other" commissions are for trading equity. Third, the market-making gains for Part IIA firms include all trading gains except those from market making in options on an exchange. I follow Stoll (1993) again and assume that 50% of the reported gains are from trading U.S. equity.

I use the following notation:

$$
\begin{aligned}
C_{\text{Total}} &= \text{Total commissions}\\
C_{\text{Listed, Exch}} &= \text{Commissions for listed equity on an exchange}\\
C_{\text{Listed, OTC}} &= \text{Commissions for listed equity traded over the counter}\\
C_{\text{Other}} &= \text{Other commissions}\\
T &= \text{Transfers between securities firms}\\
G &= \text{Market-making gains}\\
k_1 &= C_{\text{Listed, Exch}}/C_{\text{Total}}\\
k_2 &= C_{\text{Listed, OTC}}/C_{\text{Total}}\\
k_3 &= C_{\text{Other}}/C_{\text{Total}}
\end{aligned}
$$

The commissions and market-making gains for Parts II and IIA firms are:

Table A1

Commissions and Market-Making Gains from Trading U.S. Equity, in Millions of Dollars, 1980–2006

Part II firms clear trades or carry customer accounts; Part IIA firms do not. Total commission revenue is C_{Total}; $C_{\text{Listed, Exch}}$ and $C_{\text{Listed, OTC}}$ are commissions from trading listed equity on an exchange and over the counter. Other commissions, C_{Other}, for Part II firms are total commissions minus those on listed equity and listed options. For Part IIA firms, C_{Other} is total commissions minus those on listed equity traded on an exchange and on listed options. Transfers, T, are payments to other securities firms for commissions, floor brokerage, and clearance for Part II firms, and commissions paid to other brokers for Part IIA firms. Market-making gains, G, for Part II firms in 2005 and 2006 are gains from market making in OTC equities and gains on firm securities trading accounts with associated hedges. Before 2005, this revenue also includes gains on derivative trading desks in equity products. Market-making gains for Part IIA firms are all gains except those from trading options on an exchange.

	Part II Firms						Part IIA Firms				
	$C_{\text{Listed, Exch}}$	$C_{\text{Listed, OTC}}$	C_{Other}	C_{Total}	T	G	$C_{\text{Listed, Exch}}$	C_{Other}	C_{Total}	T	G
1980	4,181	97	923	5,934	633	863	504	272	843	228	213
1981	3,859	134	943	5,566	587	746	529	375	972	300	203
1982	4,336	113	959	6,163	679	789	662	464	1,208	368	232
1983	5,853	191	1,606	8,454	912	1,623	987	837	1,938	568	507
1984	4,983	172	1,315	7,306	931	1,049	932	926	1,963	653	553
1985	5,845	214	1,718	8,692	1,089	1,476	1,046	1,098	2,263	863	716
1986	7,342	301	2,243	11,076	1,380	2,065	1,315	1,427	2,900	1,165	911
1987	8,843	185	2,722	13,102	1,850	2,184	1,681	1,619	3,473	1,171	750
1988	6,162	117	2,103	9,084	1,421	1,587	1,411	1,325	2,849	867	864
1989	7,134	174	2,103	10,234	1,580	1,686	1,684	1,383	3,218	898	986
1990	5,971	203	1,984	8,836	1,454	1,537	1,613	1,435	3,196	974	900
1991	6,914	324	2,535	10,429	1,490	2,482	1,903	1,715	3,780	1,160	1,494
1992	7,533	457	3,201	11,837	1,717	2,848	2,103	2,088	4,412	1,479	2,053
1993	8,812	709	4,043	14,277	2,152	3,459	2,522	2,779	5,628	2,086	2,744
1994	8,538	637	4,154	14,105	2,548	3,263	2,582	2,867	5,742	2,105	2,399
1995	9,908	848	4,886	16,521	2,602	3,803	2,995	3,468	6,694	2,329	3,363
1996	11,073	1,010	6,630	19,674	2,997	4,350	3,269	4,683	8,191	3,493	4,585
1997	13,141	1,191	7,707	23,277	3,533	4,601	3,771	5,273	9,385	4,309	5,199
1998	14,793	1,460	8,564	26,250	4,138	5,239	4,202	5,828	10,446	5,074	4,837
1999	16,647	2,024	11,763	32,247	4,894	5,813	4,794	8,369	13,691	7,370	6,951
2000	18,054	2,370	14,339	37,020	5,356	8,188	5,786	10,646	17,087	8,946	13,449
2001	16,843	1,743	10,188	30,569	4,839	3,982	4,785	8,893	14,194	7,982	7,940
2002	17,001	2,291	10,472	31,160	4,956	1,415	4,573	8,804	13,841	8,742	6,454
2003	14,964	2,561	11,367	30,313	5,253	1,087	4,620	10,084	15,226	9,788	5,496
2004	14,860	2,945	12,798	31,992	5,327	949	4,542	10,235	15,577	10,673	5,308
2005	14,695	3,179	12,570	31,894	5,737	1,134	4,124	9,873	14,545	11,314	5,397
2006	14,901	3,036	14,269	34,018	7,924	1,769	4,272	9,894	14,845	12,286	8,034

	Part II	Part IIA
Commissions on Exchanges Trades	$C_{\text{Listed, Exch}} - k_1T$	$C_{\text{Listed, Exch}} - k_1T$
Commissions on OTC Trades	$C_{\text{Listed, OTC}} + 0.9C_{\text{Other}} - (k_2 + 0.9k_3)T$	$0.9C_{\text{Other}} - 0.9k_3T$
Market-making Gains	G	$0.5G$

E. Turnover

Most NYSE and Amex transactions are direct trades from one public customer to another. NASDAQ developed as a dealer market in which public investors sell shares to dealers who then sell them to other public investors. Thus, a transaction that transfers 100 shares from one public investor to another would typically be recorded as 100 shares traded on the NYSE and Amex, but as 200 shares traded on NASDAQ. Researchers often deal with this inconsistency by dividing reported NASDAQ volume by two. The evolution of the NASDAQ market, however, makes this rule of thumb obsolete later in the sample period. Electronic communication networks (ECNs), which allow public investors to bypass the dealer, account for a large and growing fraction of NASDAQ volume by 2001. Because ECN trades are between two public customers, there is no double counting in these transactions. Changes in the reporting rules for riskless principal transactions also reduce double counting. After 2001, a dealer who covers a client's purchase or sale with a contemporaneous trade at the same price must report the transaction as a single trade. Because of these changes, when computing the turnover in Figure 1 I divide reported NASDAQ volume by 2.0 until 2001, by 1.5 in 2002 and 2003, and by 1.25 thereafter. If I always divide NASDAQ volume by 2.0, turnover for the aggregate market in 2007 drops from 215% to 194%.

REFERENCES

Ahearne, Alan G., William L. Griever, and Francis E. Warnock, 2004, Information costs and home bias: An analysis of U.S. holdings of foreign equities, *Journal of International Economics* 62, 313–336.

Asness, Clifford, Robert Krail, and John Liew, 2001, Do hedge funds hedge? *Journal of Portfolio Management* Fall, 6–19.

Barber, Brad M., and Terrence Odean, 2000, Trading is hazardous to your wealth: The common stock investment performance of individual investors, *Journal of Finance* 55, 773–806.

Barber, Brad M., and Terrence Odean, 2001, Boys will be boys: Gender, overconfidence, and common stock investment, *Quarterly Journal of Economics* 116, 261–292.

Barber, Brad M., Terrence Odean, and Lu Zheng, 2005, Out of sight, out of mind: The effects of expenses on mutual fund flows, *Journal of Business* 78, 2095–2119.

Bogle, John C., 2008, A question so important that it should be hard to think about anything else, *Journal of Portfolio Management* 34, 95–102.

Brown, Stephen J., William N. Goetzmann, and Bing Liang, 2004, Fees on fees in funds of funds, *Journal of Investment Management* 2, 39–56.

Buffett, Warren E., 2006, *Berkshire Hathaway Inc. 2005 Annual Report*.
Fama, Eugene F., and Kenneth R. French, 2002, The equity premium, *Journal of Finance* 57, 637–659.
Fama, Eugene F., and Kenneth R. French, 2008, Mutual fund performance, Working paper, Dartmouth College.
French, Kenneth R., and James M. Poterba, 1991, Investor diversification and international equity markets, *American Economic Review* 81, 222–226.
Goetzmann, William N., and Alok Kumar, 2008, Equity portfolio diversification, *Review of Finance* (forthcoming).
Graham, John R., and Campbell Harvey, 2005, The long-run equity risk premium, *Finance Research Letters* 2, 185–194.
Karolyi, G. Andrew, and Rene Stulz, 2003, Are financial assets prices locally or globally? in George M. Constantinides, Milton Harris, and Rene Stulz, eds.: *Handbook of the Economics of Finance* (Elsevier Science B.V, Amsterdam).
Kennickell, 2003, A rolling tide: Changes in the distribution of wealth in the U.S., 1989–2001, Working paper, Securities and Exchange Commission.
Kennickell, 2006, Currents and undercurrents: Changes in the distribution of wealth, 1989–2004, Working paper, Securities and Exchange Commission.
Malkiel, Burton G., 1973, *A Random Walk Down Wall Street* (W.W. Norton & Company, New York).
Merton, Robert C., 1973, An intertemporal capital asset pricing model, *Econometrica* 41, 867–887.
Odean, Terrence, 1998, Volume, volatility, price, and profit when all traders are above average, *Journal of Finance* 53, 1887–1934.
Rea, John D., and Brian K. Reid, 1998, Trends in the ownership cost of equity mutual funds, *Investment Company Institute Perspective* 4, 1–15.
Ross, Stephen A., 1976, The arbitrage theory of capital asset pricing, *Journal of Economic Theory* 13, 341–360.
Sharpe, William F., 1991, The arithmetic of active management, *Financial Analysts Journal* 47, January/February, 7–9.
Statman, Meir, 2004, What do investors want? *Journal of Portfolio Management* 30, August, 153–161.
Statman, Meir, Steven Thorley, and Keith Vorkink, 2006, Investor overconfidence and trading volume, *Review of Financial Studies* 19, 1531–1565.
Stoll, Hans R., 1993, Equity trading costs in-the-large, *Journal of Portfolio Management* 19, Summer, 41–50.

45

September 2007

Mutual fund flows and investor returns: An empirical examination of fund investor timing ability

Geoffrey C. Friesen
University of Nebraska-Lincoln, gfriesen2@unl.edu

Travis R. A. Sapp
Iowa State University, trasapp@iastate.edu

Published in *Journal of Banking & Finance* 31:9 (September 2007), pp. 2796–2816;
doi 10.1016/j.jbankfin.2007.01.024
http://www.sciencedirect.com/science/journal/03784266

Submitted September 11, 2006; accepted January 9, 2007; published online April 29, 2007.

Mutual fund flows and investor returns: An empirical examination of fund investor timing ability

Geoffrey C. Friesen* and Travis R. A. Sapp†

* College of Business, CBA 237, University of Nebraska–Lincoln,
Lincoln, NE 68588-0490, USA (email gfriesen2@unl.edu)

† College of Business, 3362 Gerdin Business Bldg., Iowa State University,
Ames, IA 50011-1350, USA (*Corresponding author*, email trasapp@iastate.edu)

Abstract

We examine the timing ability of mutual fund investors using cash flow data at the individual fund level. Over 1991–2004, equity fund investor timing decisions reduce fund investor average returns by 1.56% annually. Underperformance due to poor timing is greater in load funds and funds with relatively large risk-adjusted returns. In particular, the magnitude of investor underperformance due to poor timing largely offsets the risk-adjusted alpha gains offered by good-performing funds. Investors in both actively managed funds and index funds exhibit poor investment timing. We demonstrate that our empirical results are consistent with investor return-chasing behavior.

Keywords: mutual fund performance, fund cash flows, investor timing, fund clienteles

1. Introduction

Mutual fund investors can enhance their returns by selecting superior funds, advantageously timing their cash flows to the fund, or both. Gruber (1996) and Zheng (1999) suggest that investors have the ability to select funds with superior subsequent performance, a result referred to as the "smart money" effect. These studies find that the short-term performance of funds experiencing positive net cash flow appears better than those experiencing negative net cash flow. Sapp and Tiwari (2004), however, demonstrate that the smart money effect is explained by stock return momentum over the short term. Further research by Frazzini and Lamont (2006) suggests that poor fund selection decisions end up costing longer-term investors (those who do not rebalance quarterly) about 0.84% per year, a result they dub the "dumb money" effect.

In this paper we focus on the second possible method by which investors may enhance their returns, which is not explicitly addressed by the above studies. We ask whether mutual fund investors make good investment decisions strictly in the timing of their cash flows. That is, for any given fund, do equity fund investors put cash in and take cash out at the right time on average? It is well established that inflows to mutual funds are strongly correlated with past fund performance (Ippolito, 1992). Less clear is the impact of investor timing decisions on investor returns. And while numerous studies have examined the timing ability of mutual fund managers or other investment professionals, ours is the first comprehensive study to examine the timing ability of mutual fund investors using cash flow data at the individual fund level.[1]

We use the dollar-weighted return, derived as the internal rate of return of money under management, to measure the performance of fund investors, and time-weighted returns to measure the performance of the fund. Because a time-weighted average return ignores month-to-month variation in assets under management, it measures the net return earned by the fund manager, or equivalently, the buy-and-hold return on a dollar invested over the entire sample period. In contrast, a dollar-weighted return explicitly accounts for net cash flows into and out of the fund over time, reflecting the average investor's performance during the sample period. We measure investor timing ability with a statistic hereafter referred to as the "performance gap", defined as the time-weighted return minus the dollar-weighted return.

The dollar-weighted return measure is particularly well-suited to the focus of this paper because dollar-weighted returns carry the implicit assumption that new cash flows are reinvested over future periods, whereas alternative measures focus only on a single period return, possibly weighting this return with current period cash flow. Specifically, other studies examining investor behavior (e.g. Zheng, 1999; Sapp and Tiwari, 2004]) impute the fund return, or alpha, to the fund investor at a single point in time. These studies implicitly assume that new money is put into the fund for one period, earns the return generated by the fund, and then is immediately taken out. These measures do not track the impact of multiple period returns on a single cash flow. In reality, the current month's positive net cash flows often remain, either in whole or in part, invested in the fund for multiple periods. Moreover, the impact of cash outflows should include not only the current month's missed return, but the opportunity cost of missed returns in future months as well. The dollar-weighted return methodology captures the interaction between all cash flows and returns to a fund over the entire sample period, thus measuring the full impact of investor cash flow timing decisions.

[1] Studies on the timing ability of fund managers include Bollen and Busse (2001), Dellva (2001), Volkman (1999), Daniel *et al.* (1997), Lee and Rahman (1990), Chang and Lewellen (1984), and Henriksson (1984). Timing ability has also been examined in the context of investment newsletter recommendations (Graham and Harvey, 1996), portfolio managers (Elton and Gruber, 1991) and investment advisors (Kleiman *et al.*, 1996; Cumby and Modest, 1987).

Prior studies have examined investors' dollar-weighted returns, but none have used cash flows at the individual fund level. For example, Nesbitt (1995) examines time-weighted and dollar-weighted returns at the aggregate level for 17 categories of mutual funds over the 1984–1994 period. He reports that, on average, investors' dollar-weighted annual returns from these categories are 1.08% less than time-weighted returns. Braverman *et al.* (2005) examine aggregate mutual fund flows and report that the annual dollar-weighted return is significantly lower than the buy-and-hold return over multiple time periods. They speculate that this finding may possibly be due to either time-varying expected returns or investor sentiment.

The use of aggregate cash flow data in these prior studies potentially biases one's inferences about investor behavior for two reasons. First, aggregation of data, and in particular of individual fund net cash flows and returns, which can be either positive or negative, discards potentially important information.[2] Second, this approach precludes any possibility of investor fund selection ability and does not afford an opportunity to examine possible differential timing performance among various fund clienteles. By using fund-level data, we are able to individually measure the timing performance of investors who choose "good" funds and investors who choose "poor" funds. Thus, the current study contributes to the literature by measuring investor timing ability while also explicitly controlling for any fund selection ability investors may possess. Our fund-level approach also has the benefit of allowing for an extensive analysis of the cross-sectional variation in investor timing performance in order to shed additional light on fund investor behavior.

For the 7,125 equity mutual funds in our sample we compute monthly dollar-weighted returns over 1991–2004 and find that the geometric average monthly return is 0.62%, while the average monthly dollar-weighted return is 0.49%. Thus, investors underperform by about 0.13% per month, or 1.56% annually, relative to a buy-and-hold strategy. This performance gap is twice as large for load funds (0.16% per month) as for no-load funds (0.08% per month). In order to distinguish between investors based on the quality of fund they choose, we compute the risk-adjusted performance, or alpha, of each fund over the sample period according to both the Fama and French (1993) 3-factor model and the Carhart (1997) 4-factor model. Using either measure, we find that poor investor timing is significantly associated with better-performing funds. More interesting yet, we find that the alpha-gain that is potentially available to investors even in good-performing funds under either benchmark measure is largely erased by the poor timing of investors in these funds. This finding is similar in spirit to the story put forth in Frazzini and Lamont (2006), where investors fail to benefit from superior performance due to entering and exiting at the wrong time.

We document further significant cross-sectional variation in the difference between time-weighted and dollar-weighted returns. The performance gap is found to be largest among the largest quintile of funds in our sample. The size of the performance gap is also

[2] To see how this could potentially impact estimation of investor timing performance, consider two funds, X and Y. Investors in Fund X display poor timing primarily through positive cash flows to the fund that occur ahead of low returns, thus generating a large measured performance gap for Fund X. Investors in Fund Y display poor timing primarily through negative cash flows that occur ahead of high returns, thus generating a large measured performance gap for Fund Y. Upon aggregating the cash flows and returns of these two funds, it is possible that no performance gap at all would be detected in the aggregate data.

increasing in fund load, turnover, and length of fund history. Overall, the evidence suggests that larger, more costly funds seem to attract less-sophisticated investors.

Analysis of fund style shows that underperformance due to timing is negatively correlated with value-style funds, but is positively associated with momentum-style funds. We find a significant performance gap for both index and non-index funds, indicating that some index fund investors are timing their investments through these low-cost vehicles, though the gap is smaller at 0.05% per month, versus 0.13% for non-index funds. We also calculate separately the dollar-weighted returns on positive and negative net cash flows for each fund. We find that on average, poorly timed purchase decisions cost investors about 0.06% per month and poorly timed withdrawals cost investors approximately 0.15% per month. We demonstrate through simulation that our empirical results are consistent with investor return-chasing behavior.

Finally, for comparison with equity funds, we examine bond funds and money market funds. We find that the average monthly performance gap over 1991–2004 is much smaller for bond funds at 0.02%, and is nearly flat for money funds at 0.004%. The poor timing phenomenon thus seems to be largely unique to equity mutual funds, suggesting either more sophisticated, or perhaps less active, investors in the bond and money funds.

Our study adds to the growing literature on the behavior and performance of mutual fund investors. By analyzing investor timing at the individual fund level, our methodology preserves cross-sectional differences in the timing performance of investors in individual funds. We not only show that attempts to time the market by fund investors are on average detrimental to investor returns, but we shed light on which fund investors are most likely to exhibit poor timing. Our results are consistent with investor return-chasing behavior. However, it is sobering to reiterate that the performance gap due to poor investor timing largely offsets the value added by actively managed funds in terms of alpha for the subset of funds that does indeed offer a positive alpha. Hence, even investors who select the best funds on average sacrifice the potential benefit due to poor timing of cash flows. Overall, our results commend the relative appeal of a simple "buy-and-hold" strategy to the average investor.

The rest of the paper is organized as follows. Section 2 describes the data and outlines our return measurement and performance benchmarking methodology. Section 3 presents the empirical results on investor timing performance and examines the relationship between fund characteristics and the timing performance gap. Section 4 explores possible explanations for investor return behavior, and Section 5 concludes.

2. Data and return measurement methodology

2.1. Sample description

Our sample is taken from the CRSP Survivor-Bias Free US Mutual Fund Database, and includes all domestic common stock funds that exist at any time during the period 1991–2004 for which monthly total net assets (TNA) values exist. Funds with fewer than 12 monthly observations are excluded from the sample. We also exclude international, sector, balanced, and specialized funds, as the benchmarking models employed in our cross-sectional analysis may be inappropriate for these funds. Monthly returns are adjusted to account for multiple fund distributions on the same day, as suggested by Elton *et al.* (2001).

Since the dollar-weighted return is an internal rate of return measure, it suffers from the multiple solutions problem when monthly fund cash flows repeatedly change sign. However, many of these solutions are either complex numbers or real numbers that are less than −100%. For the vast majority of funds, there exists only one real root greater than −100%. Due to the limited liability constraint inherent in a mutual fund investment, we retain only funds with a unique dollar-weighted return above −100%, which yields a sample of 7125 funds. Unless otherwise noted, all of our analysis is conducted for these 7125 funds. We also note that, for purposes of this study, fund share classes are treated as distinct funds.

Table 1 reports descriptive statistics for the fund sample. The average fund has nearly half a billion dollars under management and experiences monthly net cash flows of 0.65% of TNA. We also note that average annual fund turnover is 92% of fund assets, the average total load fee is 2.32%, and the average annual fund expense ratio is 1.42%.

2.2. *Measurement of returns and cash flows*

Denote the return for fund j in month t to be r_{jt}. The geometric average monthly return for fund j is calculated as

$$\bar{r}_j^g = \left(\prod_{t=1}^{T}(1+r_{jt})\right)^{1/T} - 1 \tag{1}$$

Table 1. Sample statistics

	Mean	Median	25th percentile	75th percentile	Standard deviation
Total net assets ($ millions)	478.94	466.33	390.85	545.47	101.84
Monthly net cash flow ($ millions)	3.09	2.97	1.73	4.66	1.70
Turnover (%/year)	91.67%	88.44%	85.47%	96.11%	12.58%
Maximum front-end load fee (%)	1.51%	1.27%	1.22%	1.70%	0.47%
Maximum total load fee (%)	2.32%	2.26%	2.19%	2.30%	0.20%
Expense ratio (%/year)	1.42%	1.44%	1.36%	1.47%	0.12%

The table presents summary statistics on the mutual fund sample obtained from the CRSP Survivor-Bias Free US Mutual Fund Database. The sample includes all US equity mutual funds that existed at any time during January 1991 through December 2004 for which monthly total net assets (TNA) values exist. Sector funds, international funds, balanced funds and specialized funds are excluded. The final sample contains 7125 funds. The monthly net cash flow for fund j in month t is $\text{NCF}_{j,t} = \text{TNA}_{j,t} - \text{TNA}_{j,t-1}(1 + r_{j,t})$, where, $\text{NCF}_{j,t}$ denotes the monthly net cash flow for fund j in month t, $\text{TNA}_{j,t}$ is the total net assets for fund j at the end of month t, and $r_{j,t}$ is the fund's return in month t. Turnover is defined as the minimum of aggregate purchases or sales of securities during the year, divided by the average TNA. Maximum front-end load is the maximum percent charges applied at the time of purchase, while maximum total load fees equals maximum front-end load fees plus maximum sales charges paid when withdrawing money from the fund. The expense ratio is the percentage of total investment that shareholders pay for the fund's operating expenses. For each item, we compute the cross-sectional averages in each year from 1991 to 2004. The reported statistics are computed from the time-series of the 14 annual cross-sectional averages for each item.

Geometric returns are appropriate measures of past fund manager performance, and also measure the average return on a dollar invested during the entire sample period. The dollar-weighted average return measures the return weighted by the amount of money invested at each point in time, and thus captures the average return earned by fund investors. The dollar-weighted average monthly return for fund j is defined as the rate of return at which the accumulated value of the initial TNA, plus the accumulated value of net cash flows, equals the actual TNA at the end of the sample period:

$$\bar{r}_j^{\mathrm{dw}} : \mathrm{TNA}_0(1+\bar{r}_j^{\mathrm{dw}})^T + \sum_{t=1}^{T} \mathrm{NCF}_t(1+\bar{r}_j^{\mathrm{dw}})^{(T-t)} = \mathrm{TNA}_T \tag{2}$$

where

$$\mathrm{NCF}_{j,t} = \mathrm{TNA}_{j,t} - \mathrm{TNA}_{j,t-1}(1 + r_{j,t}) \tag{3}$$

Here, $\mathrm{NCF}_{j,t}$ denotes the monthly net cash flow for fund j in month t and $\mathrm{TNA}_{j,t}$ is the total net assets for fund j at the end of month t.[3] All investor cash flows are implicitly assumed to occur discretely at the end of each month. We follow Gruber (1996) and assume that investors in merged funds place their money in the surviving fund and continue to earn the return on the surviving fund. Because the holdings of the investor are identical to the holdings of the fund itself at any point in time, no risk adjustment is necessary in order to measure investor timing. Our measure of investor timing for fund j, which we refer to as the performance gap, is computed by subtracting the dollar-weighted return in Equation (2) from the geometric fund return in Equation (1):

$$\text{Performance gap}_j = \bar{r}_j^{\mathrm{g}} - \bar{r}_j^{\mathrm{dw}} \tag{4}$$

We do note the possibility that some sophisticated investors may shift their portfolio holdings among other asset classes as part of an overall asset allocation or risk-reduction strategy. Any possible effect on overall investor performance from such activity will not be captured by this measure. This timing performance measure simply judges the success of investor cash flows against a buy-and-hold strategy in the respective fund.

2.3. *Measurement of fund performance*

For our cross-sectional analysis of investor timing ability, we wish to classify funds according to their risk-adjusted performance. We evaluate fund performance using two commonly employed benchmark models: the Fama and French (1993) 3-factor model, and a 4-factor model as in Carhart (1997). Specifically, the Fama–French 3-factor model is given by:

$$r_{p,t} = a_p + \beta_{1,p}\mathrm{RMRF}_t + \beta_{2,p}\mathrm{SMB}_t + \beta_{3,p}\mathrm{HML}_t + e_{p,t} \tag{5}$$

Here, $r_{p,t}$ is the monthly return on fund p in excess of the one month T-bill return; RMRF is the excess return on a value-weighted market portfolio; and SMB and HML are returns on zero-investment factor-mimicking portfolios for size and book-to-market. The Carhart

[3] While this framework for calculating net cash flows is standard in the literature, we also confirm that our results are robust to assuming that cash flows occur at the beginning or middle of the month. Ber and Ruenzi (2006) study the general suitability of using net imputed cash flows as opposed to actual inflows and outflows and conclude that the net cash flow measure serves as an appropriate and unbiased measure.

4-factor benchmarking model is given by:

$$r_{p,t} = a_p + \beta_{1,p}\,\mathrm{RMRF}_t + \beta_{2,p}\,\mathrm{SMB}_t + \beta_{3,p}\,\mathrm{HML}_t + \beta_{4,p}\,\mathrm{UMD}_t + e_{p,t} \tag{6}$$

where $r_{p,t}$, RMRF, SMB, and HML are as in the Fama–French 3-factor model, and UMD is the return on the zero-investment factor-mimicking portfolio for one-year momentum in stock returns. For each model, alpha is computed for each fund from all available return data over the sample period, with a minimum of 12 return observations being required for estimation.

3. Empirical results

3.1. *Investor timing performance*

We compute arithmetic, geometric and dollar-weighted average returns for each fund in our sample, and the results are reported in Panel A of Table 2. For the average fund, investors earn 0.13% less per month (1.56% annually) than the fund itself. For the median fund, the monthly performance gap is 0.11% (1.32% annually). Panels B and C report average returns for index and non-index funds, respectively. Interestingly, investors in passively managed funds appear to attempt market timing, though we note that the mean monthly performance gap of 0.13% for actively managed funds is larger than the gap of 0.05% observed for index funds. Panels D and E report average returns for load and no-load funds, respectively. The monthly performance gap of 0.08% for no-load funds is about half the gap of 0.16% observed for load funds. Taken together, Panels A–E in Table 2 suggest that mutual fund investors on average underperform their chosen funds by between 1% and 2% per year due to the timing of their cash flows.

In order to see whether differences in investor timing ability exist between fund objective categories, we sort funds based on their CRSP SI-Objective variable and report summary statistics for each objective category in Table 3. We find that the performance gap is positive and significant for all six major objective categories, although growth-oriented categories in general have the largest performance gaps while income-oriented funds have the smallest. The largest performance gap is seen for aggressive growth funds at 0.25% per month (3.00% annually), and this category also exhibits the largest cross-sectional variability in fund performance.

A potential concern is that our results may be driven by small funds with relatively fewer assets under management, since each fund receives an equal weight in the reported average regardless of size. If true, then our reported average performance gap need not represent the performance gap for the average dollar invested in equity funds. However, we find that our results are in fact driven by the larger funds in our sample. In Table 4 we report the performance gap when funds are sorted into quintiles based on total assets. The table reveals a monotonic relationship between the performance gap and fund size categories, where underperformance is the largest for the largest quintile of funds. The timing performance gap is significantly positive for all size categories except for the smallest funds, where it is indistinguishable from zero. The largest quintile of funds has an average monthly performance gap of 0.19% (2.28% annually).This suggests that a simple average of all funds may actually understate the performance gap on the average dollar invested in equity funds, since the performance gap is greatest among the largest funds.

Table 2. Fund returns and investor timing performance

	Mean	Median	25th percentile	75th percentile	Standard deviation
Panel A: All funds (n = 7,125)					
Arithmetic monthly return	0.74	0.83	0.39	1.21	0.91
Geometric monthly return	0.62	0.69	0.24	0.11	0.96
Dollar-weighted monthly return	0.49	0.62	0.02	1.07	1.02
Performance gap	0.13	0.11	−0.11	0.35	0.53
(t-stat.)	(20.70)				
Panel B: Index funds (n = 416)					
Arithmetic monthly return	0.73	0.77	0.33	1.07	0.84
Geometric monthly return	0.62	0.65	0.20	0.95	0.87
Dollar-weighted monthly return	0.57	0.62	0.13	1.07	0.91
Performance gap	0.05	0.06	−0.21	0.28	0.45
(t-stat.)	(2.27)				
Panel C: Non-index funds (n = 6,709)					
Arithmetic monthly return	0.74	0.82	0.39	1.18	0.89
Geometric monthly return	0.60	0.68	0.24	1.06	0.94
Dollar-weighted monthly return	0.47	0.59	0.01	1.04	1.01
Performance gap	0.13	0.11	−0.10	0.36	0.54
(t-stat.)	(19.72)				
Panel D: Load funds (n = 4,408)					
Arithmetic monthly return	0.68	0.76	0.34	1.11	0.81
Geometric monthly return	0.53	0.63	0.19	0.98	0.84
Dollar-weighted monthly return	0.38	0.50	−0.07	0.95	0.94
Performance gap	0.16	0.12	−0.09	0.37	0.51
(t-stat.)	(20.83)				
Panel E: No-load funds (n = 2,717)					
Arithmetic monthly return	0.85	0.90	0.48	1.31	0.98
Geometric monthly return	0.70	0.76	0.34	1.20	1.05
Dollar-weighted monthly return	0.63	0.73	0.20	1.20	1.09
Performance gap	0.08	0.07	−0.14	0.33	0.60
(t-stat.)	(6.95)				

For each fund, we calculate the average monthly arithmetic, geometric and dollar-weighted returns over the entire sample period. Performance gap is the difference between fund geometric and dollar-weighted returns. Panel A reports statistics on the full sample of funds. Panel B reports returns separately for index funds, while Panel C reports returns for non-index funds. Panel D reports returns for load funds, and Panel E reports returns for no-load funds. t-Statistics for the mean performance gap are reported in parentheses. Returns are percent per month.

3.2. *Fund alphas and the timing performance gap*

By measuring investor timing ability at the individual fund level we are able to examine whether there is any apparent relationship between timing performance and the quality of the fund selected by an investor. For this purpose we compute a risk-adjusted return, or alpha, according to both the Fama and French (1993) 3-factor and Carhart (1997) 4-factor benchmark models for each fund over the sample period. Using this measure of fund quality, we then sort all funds into deciles based on the alpha measure of fund performance.

Table 3. Timing performance by fund objective

	Mean values			Median values				
	Geometric monthly return	Dollar-weighted monthly return	Performance gap	Geometric monthly return	Dollar-weighted monthly return	Performance gap	Std. dev. of avg. fund return	Std. dev. of performance gap
Aggressive growth $N = 456$	0.37	0.13	0.25 (7.52)	0.58	0.34	0.21	1.31	0.71
Small-cap growth $N = 1428$	0.91	0.75	0.16 (10.25)	0.90	0.79	0.14	1.00	0.59
Mid-cap growth $N = 825$	0.77	0.64	0.13 (6.79)	0.82	0.75	0.09	0.94	0.55
Growth $N = 2509$	0.40	0.26	0.14 (13.46)	0.52	0.42	0.11	0.91	0.52
Growth and income $N = 1570$	0.57	0.51	0.06 (5.32)	0.63	0.56	0.06	0.70	0.45
Income-growth $N = 319$	0.71	0.68	0.03 (1.98)	0.72	0.76	0.02	0.51	0.27

For each fund, we calculate the average monthly arithmetic, geometric and dollar-weighted returns over the entire sample period. Performance gap is the difference between fund geometric and dollar-weighted returns. Funds are divided into objective categories using the CRSP SI-Objective variable, and summary statistics are reported for each objective category. Standard deviations are reported for the average geometric return and performance gap. *t*-Statistics for the mean performance gap are reported in parentheses. Returns are percent per month.

Table 4. Timing performance by fund size

	(small) Quintile 1	Quintile 2	Quintile 3	Quintile 4	(large) Quintile 5
Average TNA (millions)	1.30	8.56	30.70	100.79	1251.65
Arithmetic return	0.58	0.65	0.74	0.84	0.91
Geometric return	0.44	0.51	0.59	0.69	0.76
Dollar-weighted return	0.43	0.39	0.45	0.52	0.57
Performance gap	0.01	0.12	0.14	0.17	0.19
(t-stat.)	(0.77)	(8.69)	(11.42)	(13.29)	(17.67)

For each fund, we calculate the average monthly arithmetic, geometric and dollar-weighted returns over the entire sample period. Performance gap is the difference between fund geometric and dollar-weighted returns. Funds are divided into quintiles based upon average total net assets (TNA). The cross-sectional averages for each TNA-based quintile are reported. Quintile 1 contains the smallest funds and quintile 5 contains the largest funds. Returns are percent per month.

Panel A of Table 5 sorts funds by 3-factor alpha and reports the mean 3-factor alpha and performance gap for each decile. We first of all note that the timing performance gap is positive and significant for all deciles of alpha-sorted funds. The average annual 3-factor alpha for all funds in the sample is −0.18% per month (−2.18% annually), and only the top three deciles of funds have an average alpha that is positive. The relationship between investor timing underperformance and the risk-adjusted performance of the fund is quite strong, with a Spearman rank correlation of 0.84, significant at the 1% level. For the decile of best performing funds, the 3-factor alpha is an impressive 0.57% per month, but this subset of funds also has the largest performance gap at 0.38% per month due to poor cash flow timing by investors. We also separately report the average alpha and average performance gap for the subset of 1902 funds that has a positive alpha. It is interesting to note that the alpha-gains of 0.27% per month offered by these good-performing funds is largely offset by average investor underperformance of 0.25% per month due to poor timing decisions.[4]

Panel B of Table 5 ranks funds into deciles according to the 4-factor alpha performance measure and reports the mean 4-factor alpha and performance gap for each decile of funds. Controlling for stock return momentum has no material effect on the results, which are nearly indistinguishable from those of the 3-factor analysis in Panel A. Investors in the better-performing funds again exhibit the poorest cash flow timing, which to a large extent offsets the superior performance offered by these funds. For the 1918 funds that generate a positive alpha, the potential gain of 0.23% per month is only slightly larger than the average investor underperformance of 0.18% per month due to poor cash flow timing.

3.3. Determinants of the performance gap

We have conducted several univariate sorts of the data which have revealed some interesting features of investor timing underperformance. We now analyze the determinants of the performance gap controlling for a number of fund characteristics such as fund age, size, expenses, load, turnover, level of cash flow, volatility, and a measure of overall performance. For each fund, the mean level of each fund characteristic over the sample period is employed. Model I in Table 6 includes among the regressors the mean return of the fund

[4] Note that alpha is computed as an arithmetic return whereas timing underperformance is computed from geometric returns. Therefore, a comparison of the two measures is only approximate and suggestive.

Table 5. Timing performance for deciles formed on fund alpha

3-Factor alpha performance decile	3-Factor alpha	Timing performance gap	Timing performance gap *t*-stat.
Panel A: Performance ranked on 3-factor alpha			
1 Worst	−0.993	0.068	2.42
2	−0.512	0.080	4.65
3	−0.369	0.054	3.34
4	−0.277	0.036	2.29
5	−0.201	0.076	5.02
6	−0.131	0.094	5.92
7	−0.061	0.146	8.98
8	0.015	0.171	9.04
9	0.139	0.166	7.99
10 Best	0.571	0.378	14.19
All funds	−0.182	0.127	
Alpha > 0 funds (*N* = 1,902)	0.273	0.252	
Spearman rank correlation	0.84 ***		

4-Factor alpha performance decile	4-Factor alpha	Timing performance gap	Timing performance gap *t*-stat.
Panel B: Performance ranked on 4-factor alpha			
1 Worst	−0.971	0.110	3.98
2	−0.520	0.092	4.70
3	−0.369	0.118	6.11
4	−0.273	0.079	4.59
5	−0.197	0.118	7.06
6	−0.129	0.120	7.05
7	−0.061	0.103	6.56
8	0.017	0.136	8.42
9	0.120	0.164	8.36
10 Best	0.487	0.228	8.87
All funds	−0.190	0.127	
Alpha > 0 funds (*N* = 1,918)	0.233	0.182	
Spearman rank correlation	0.76 **		

Panel A reports the mean alpha and mean performance gap for deciles of funds sorted on 3-factor alpha. Panel B reports the mean alpha and mean performance gap for deciles of funds sorted on 4-factor alpha. Three-factor and 4-factor alphas are estimated for each fund according to Equations (5) and (6), respectively, in the text using all available fund returns in the sample period. All returns are percent per month.
** Significant at the 5% level. *** Significant at the 1% level.

over the sample period as a measure of performance. Model II replaces the raw return with the fund 3-factor alpha as a measure of performance, and Models III and IV adopt the 4-factor alpha as a performance measure. Model IV also includes the estimated factor loadings for size, book-to-market, and momentum in order to control for fund style.

Results show that the size of the performance gap is increasing in fund load fees, turnover, and length of return history, although the significance of turnover is generally marginal. The positive relation between timing underperformance and fund turnover is particularly intriguing, since both dollar-weighted and geometric returns are measured net of expenses and trading costs. The evidence indicates that older and more expensive funds are associated with an investor clientele that is especially poor at cash flow timing. Fund

Table 6. Determinants of the performance gap

	Model I	Model II	Model III	Model IV
Intercept	−0.446	−0.118	−0.152	−0.104
	(−9.95)	(−2.69)	(−3.45)	(−2.36)
Number of returns	0.001	0.001	0.001	0.001
	(5.54)	(4.50)	(4.93)	(5.25)
Average TNA	0.001	0.001	0.001	0.000
	(0.54)	(0.48)	(0.53)	(0.05)
Average fund expenses	3.167	2.133	1.607	1.504
	(1.31)	(0.83)	(0.61)	(0.56)
Average total load	0.978	0.785	0.824	0.782
	(2.85)	(2.24)	(2.31)	(2.18)
Average turnover	0.009	0.011	0.011	0.009
	(1.94)	(2.30)	(2.29)	(1.86)
Average net cash flow (% of TNA)	−0.001	−0.001	−0.001	−0.001
	(−0.62)	(−0.86)	(−0.83)	(−0.80)
Average return	13.387			
	(11.03)			
Standard deviation of returns	7.094	1.143	1.750	0.604
	(10.01)	(1.40)	(2.07)	(0.74)
3-Factor alpha		13.690		
		(4.36)		
4-Factor alpha			9.978	8.993
			(3.19)	(2.70)
SMB factor loading				−0.054
				(−1.80)
HML factor loading				−0.083
				(−2.89)
UMD factor loading				0.118
				(1.99)
3-Factor tracking error		0.050		
		(4.22)		
4-Factor tracking error			0.054	0.064
			(4.03)	(4.52)
Adj. R^2	0.083	0.059	0.050	0.057

For each equity mutual fund, we calculate the difference between geometric and dollar-weighted returns, which we label the fund's performance gap. The performance gap is the dependent variable in a linear regression on the fund characteristics listed in the first column of the table. 3-Factor and 4-factor alphas are estimated for each fund according to Equations (5) and (6), respectively, in the text using all available fund returns in the sample period. The regression coefficients are reported with White heteroskedasticity-consistent t-statistics in parentheses.

volatility, especially non-market volatility, is seen to be positively correlated with timing underperformance. We later discuss how investor return-chasing behavior can explain this finding. We note that neither fund size nor average net cash flow are significant predictors of timing performance after controlling for other fund characteristics. The fact that the level of fund net cash flows has no marginal explanatory power for the performance gap suggests that the overall rate of non-investment growth of the fund is irrelevant to investor timing performance. Load funds are typically purchased with the help of a broker or investment advisor, and our evidence suggests that those investors who are most likely relying on advice from a broker perform especially poorly from a timing standpoint. This is consistent with Bergstresser *et al.* (2006), who find that brokers typically fail to deliver any tangible benefits to their clientele.

Table 6 also confirms that the performance gap is greatest in funds with the best performance, whether measured by raw returns or by either the 3-factor or 4-factor benchmark. This is an interesting finding, because it tells us that there is no necessary connection between being able to select good funds and timing investment cash flows well. In fact the evidence is quite the opposite: investors who select a good fund are nevertheless plagued by particularly poor timing of their cash flows. Finally, we note that a size-based fund style is not correlated with the performance gap, although underperformance due to timing is negatively correlated with value-style funds and is positively associated with momentum-style funds. The finding that investors poorly time cash flows into momentum-style funds is consistent with return-chasing behavior, an issue we further explore below.

3.4. *An alternative measure of investor timing ability*

To further examine the source of timing underperformance exhibited by investors, we separately calculate the dollar-weighted returns on positive and negative cash flows to each fund. Using this approach we are able to separate the effect of net purchase and withdrawal decisions in order to determine whether these have a differential impact on investor timing performance. An additional feature of this approach is that fund total assets are ignored and investor dollar-weighted returns are therefore unaffected by changes in fund size that are due to fund returns. Thus, calculating investor returns separately for positive and negative cash flows also serves as a robustness check on our earlier results. On the other hand, care must be taken in handling these returns because the dollar amounts of positive and negative cash flow from which they are respectively derived may differ substantially.

Note that investors with positive timing ability will systematically invest more money prior to high return periods, producing a dollar-weighted return on positive net cash flows that exceeds the geometric average return. They will also systematically withdraw funds prior to low return periods, generating a dollar-weighted return on negative net cash flows that is less than the fund's geometric average return. From the investor's perspective, high dollar-weighted returns are desirable for positive cash flows, while low dollar-weighted returns are desirable for negative cash flows. In particular, the average return generated by the fund serves as the relevant benchmark against which we compare the average investor returns on positive and negative cash flows.

Define, $\mathrm{NCF}^{+}_{j,t} \equiv \max(\mathrm{NCF}_{j,t}, 0)$ and $\mathrm{NCF}^{-}_{j,t} \equiv \min(\mathrm{NCF}_{j,t}, 0)$. The dollar-weighted return on positive net cash flows only, $\bar{r}_j^{\mathrm{dw},+}$, is defined as

$$r_j^{\mathrm{dw},+} : \sum_{t=1}^{T} \mathrm{NCF}^{+}_{j,t}(1 + r_j^{\mathrm{dw},+})^{(T-t)} = \sum_{t=1}^{T} \left(\mathrm{NCF}^{+}_{j,t} \prod_{s=t+1}^{T} (1 + r_{j,s}) \right) \tag{7}$$

and the dollar-weighted return on negative net cash flows, $\bar{r}_j^{\mathrm{dw},-}$, is defined as

$$\bar{r}_j^{\mathrm{dw},-} : \sum_{t=1}^{T} \mathrm{NCF}_{j,t}(1 + \bar{r}_j^{\mathrm{dw},-})^{(T-t)} = \sum_{t=1}^{T} \left(\mathrm{NCF}_{j,t} \prod_{s=t+1}^{T} (1 + r_{j,s}) \right) \tag{8}$$

Table 7 reports the dollar-weighted return calculated separately for positive and negative net cash flows, according to Equations (7) and (8). We find that the dollar-weighted return on positive net cash flows is 0.56% per month for the average fund, while the dol-

Table 7. Investor returns by signed cash flow

	Mean	Median	25th percentile	75th percentile	Standard deviation
Dollar-weighted return on positive net cash flows	0.56	0.63	0.11	1.12	1.04
Dollar-weighted return on negative net cash flows	0.77	0.80	0.29	1.40	1.35
Difference (positive − negative)	−0.21	−0.14	−0.53	0.13	0.95
(t-stat.)	(19.55)				

lar-weighted average return on negative net cash flows is 0.77%. Thus the average new dollar invested earned 0.56% per month, while the average dollar withdrawn would have earned 0.77% had it remained in the fund, representing an unfavorable overall difference of 0.21% in monthly return. Moreover, comparing each to the average fund return of 0.62%, we see that poorly timed purchase decisions cost investors about 0.06% per month and poorly timed withdrawals cost investors approximately 0.15% per month. Overall, poor investor withdrawal decisions hurt investors more than poor purchase decisions, though both clearly play a role in investor underperformance.

This table reports statistics on fund dollar-weighted average monthly returns computed separately on positive and negative net cash flows for the full sample of equity funds. Statistics for the difference in positive and negative cash flow returns is also reported. These returns are calculated according to Equations (7) and (8) in the text. Returns are percent per month.

One possible explanation of these results is that investors respond to poor fund performance by withdrawing assets, behaving in a manner consistent with the limits-of-arbitrage story of Shleifer and Vishny (1997). In their model, investors withdraw money after negative returns, thereby irrationally selling assets that are in fact undervalued. It is also possible that investor withdrawals are liquidity motivated, and that investor liquidity needs are most acute in periods where fund returns are poor. In either case, we can state that investors systematically withdraw funds prior to relatively good performance, and these withdrawals reduce investor returns.

3.5. *Timing performance by year*

In order to shed light on whether investor timing performance differs by time period or is sensitive to the length of fund return history, we estimate and report fund geometric and investor dollar-weighted returns on a calendar year basis. Only funds having 12 monthly returns in a given year are included in the sample for the year, and returns for the fund are computed only based on the 12 months of data for the year. Results are reported in Table 8 for each year of the sample as well as for all fund-years.

Since the mutual fund industry was growing throughout the 1991–2004 sample period, the least number of funds (296) appears in 1991, and the greatest number of funds (3765) appears in 2004. Of the 14 years in our sample, the performance gap is positive and significant in all but three—1995, 1997, and 2003—which presents no discernable pattern over time. Thus, it appears that investors tend to underperform a buy-and-hold strategy in all manner of market conditions. We also note that the dollar-weighted return on negative cash flows exceeds the dollar-weighted return on positive cash flows in all years ex-

Table 8. Yearly returns and timing performance

Year	All	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004
Number of fund-years	28,244	296	561	708	975	1182	1377	1701	2253	2512	2556	3113	3513	3732	3765
Arithmetic return	0.86	2.78	1.15	1.01	0.06	2.21	1.50	1.68	1.23	2.30	0.73	−0.40	−1.63	1.87	1.10
Geometric return	0.71	2.67	1.08	0.97	0.00	2.18	1.43	1.57	0.97	2.16	0.40	−0.65	−1.81	1.78	1.05
Dollar-weighted return	0.57	2.65	0.73	0.90	−0.20	2.21	1.40	1.66	0.93	1.88	−0.35	−1.08	−2.24	2.41	0.88
Performance gap	0.15	0.02	0.35	0.07	0.21	−0.03	0.03	−0.09	0.04	0.28	0.75	0.43	0.43	−0.62	0.17
(*t*-stat.)	(41.03)	(2.87)	(17.08)	(9.85)	(17.09)	(−4.13)	(3.01)	(−9.59)	(4.06)	(18.44)	(45.14)	(57.12)	(77.23)	(−130.6)	(40.12)
Positive NCF dollar-weighted return	0.85	2.46	1.32	0.98	−0.15	1.89	1.32	1.32	1.39	3.01	−0.69	−0.42	−2.30	2.92	1.38
Negative NCF dollar-weighted return	1.35	2.49	1.77	1.19	−0.07	2.06	1.57	1.58	2.81	3.16	−0.70	0.71	−1.46	3.10	1.89
Positive − negative	−0.50	−0.03	−0.45	−0.21	−0.08	−0.17	−0.25	−0.27	−1.42	−0.14	0.01	−1.13	−0.84	−0.18	−0.52
(*t*-stat.)	(−42.51)	(−0.39)	(−7.99)	(−7.07)	(−3.36)	(−6.22)	(−8.44)	(−8.19)	(−20.9)	(−2.29)	(0.21)	(−27.6)	(−30.2)	(−15.06)	(−23.81)

For each calendar year, we calculate arithmetic, geometric, and dollar-weighted mean monthly returns for each equity fund and report the cross-sectional average. Performance gap is the difference between the geometric and dollar-weighted returns. To be included in the sample for a particular year, a fund must have 12 monthly return observations in that year. Returns are percent per month.

cept 2000, and in that year the difference is not significantly different from zero. This indicates that poor cash withdrawal decisions are consistently more detrimental to investor timing performance than poorly timed cash flows into funds.

In the regression analysis in Table 6 we noted that a longer return history is associated with a larger performance gap. The fact that the performance gap is positive in most years, including the early years of the sample, explains why a longer return history is correlated with a larger performance gap: there is simply more time to accumulate losses. This implies that funds with longer return histories will generally have a greater influence on reported sample means. Furthermore, the yearly computation of returns allows us to gauge the extent of any possible effect. Specifically, by pooling all 28,244 fund-years of data computed by this method we obtain an average monthly performance gap of 0.15%, which is nearly identical to our previously reported value of 0.13% when using all available data for each fund. To summarize, whether we compute and weight returns by fund or by fund-year, we find that the overall performance gap is approximately the same.

4. Exploring investor behavior

4.1. Return-chasing and the performance gap

The empirical finding of a pervasive timing performance gap is consistent with a behavioral explanation where fund investors simply chase large recent returns and flee from low recent returns. Timing underperformance may occur even if investors are able to identify funds that on average outperform their peers and is likely due to a combination of the weak persistence in fund returns and investor failure to rebalance at the right time. Intuitively speaking, if fund returns are serially uncorrelated and investors buy in following returns far above the mean while taking cash out following returns far below the mean, they will on average lose due to the tendency of outcomes to cluster at the mean. Even in the presence of some weak return persistence, investors may over-estimate their ability to exploit this persistence. If active investors do not rebalance at the right times, they can still suffer inferior performance due to poor cash flow timing.

Numerous studies in the experimental psychology literature demonstrate that individual cognitive biases are often state dependent. For example, overconfidence tends to be most pronounced in situations where information is ambiguous and predictability is low (Griffn and Tversky, 1992) and the task is of moderate to extreme difficulty (Fischoff *et al.*, 1982). Overconfident investors over-estimate the precision of their information, trade too frequently, and as a result experience poor investment performance (Odean, 1998).

Kahneman and Tversky (1972) examine the representativeness heuristic, defined as a subjective judgment of the extent to which an event is similar in essential properties to the parent population. They demonstrate that individuals often over-estimate the degree to which a single event is similar to the parent population. Mutual fund investors who exhibit the representativeness heuristic will over-estimate the predictability of fund returns, believing that a single large return is indicative of a fund with a high mean return. This could lead to return-chasing behavior and generate a performance gap between investor returns and the returns of the underlying fund.

Evidence on fund return volatility and fund style is consistent with the behavioral hypothesis just described. Specifically, the results reported in Table 3 on the volatility of average returns for the funds sorted by investment objective show a strong positive correlation between return volatility and the performance gap. We also note that in the regressions in Table 6, fund total volatility and tracking error are both significant predictors of timing underperformance. Regarding fund style, in Table 6 timing underperformance was found to be significantly correlated with momentum style funds. Momentum style funds have greater recent return persistence than other funds and may serve as a stronger inducement for investors to attempt timing. While consistent with a behavioral explanation, we note that this evidence is only suggestive. To explore the issue further, we next present simulation evidence that is also suggestive of a return-chasing explanation for timing underperformance.

4.2. Simulation evidence

In this section, we use simulated data to study how our measures of performance vary for different specifications of investor behavior. We simulate a sample of 7,125 funds with 36 monthly return observations each. Monthly returns are calibrated to correspond to the average return in our actual data sample and are assumed to be independent draws from a normal distribution with mean return 0.75% and standard deviation of 5%. Net cash flows are assumed to occur at the end of each month. The net cash flow, as a percentage of the end-of-month TNA (after returns), is determined by one of five specifications. All specifications consist of a random liquidity component for fund j in month t, $\varepsilon_{jt} \sim N(0, s.d. = 1\%)$. In addition, specifications 2–5 consist of a behavioral component. The sensitivity of cash flows to returns employed in each model was calibrated through regression using the mutual fund sample.[5] Let

$$I^{+} = \begin{cases} 1 & \text{if } r_{jt-1} > 0.75\%, \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad I^{-} = \begin{cases} 1 & \text{if } r_{jt-1} \leq 0.75\% \\ 0 & \text{otherwise} \end{cases}$$

$$\text{Specification 1:} \quad \text{NCF}_{jt} = \varepsilon_{jt} \tag{9a}$$

$$\text{Specification 2:} \quad \text{NCF}_{jt} = \varepsilon_{jt} + 3(r_{jt-1} - 0.0075) \tag{9b}$$

$$\text{Specification 3:} \quad \text{NCF}_{jt} = \varepsilon_{jt} + 3(r_{jt-1} - 0.0075) \cdot I^{+} \tag{9c}$$

$$\text{Specification 4:} \quad \text{NCF}_{jt} = \varepsilon_{jt} + 3(r_{jt-1} - 0.0075) \cdot I^{-} \tag{9d}$$

$$\text{Specification 5:} \quad \text{NCF}_{jt} = \varepsilon_{jt} - 3(r_{jt-1} - 0.0075) \tag{9e}$$

Under specification 1, net cash flows are random. End-of-month net cash flows are correlated with the prior month's return under specification 2, so that investors make positive investments in funds with above average returns, and withdraw money from funds experiencing below average returns. The magnitude of the cash flow is directly proportional to the difference between the actual return and the average return. In specification 3, investors chase hot funds, but net cash flows are random for funds with poor returns. Investors flee from poor performers under Specification 4, but cash flows to hot funds are random. Specification 5 simulates a contrarian strategy, where investors sell funds af-

[5] For each fund, we regress percentage net cash flow on lagged mean-centered returns. The average cross-sectional coefficient from these regressions is 3.06, thus motivating our choice of 3.0 for the performance-cash flow sensitivity coefficient on lagged returns in (9b) through (9e).

Table 9. Returns and performance gaps for simulated return data

	Actual data	Simulated data				
		Scenario 1	Scenario 2	Scenario 3	Scenario 4	Scenario 5
Geometric average return	0.62	0.63	0.63	0.63	0.63	0.63
Dollar-weighted return	0.49	0.63	0.28	0.47	0.48	1.02%
Performance gap	0.13	−0.00	0.35	0.16	0.15	−0.39
(t-stat.)	(20.70)	(−0.93)	(94.07)	(52.91)	(21.25)	(−98.50)
Dollar-weighted return on positive net cash flows	0.56	0.60	0.41	0.40	0.52	0.82
Dollar-weighted return on negative net cash flows	0.77	0.60	0.61	0.59	0.66	0.60
Difference (positive − negative)	−0.21	−0.00	−0.20	−0.19	−0.13	0.22
(t-stat.)	(−19.55)	(−0.35)	(−81.32)	(−51.18)	(−41.68)	(100.28)

Returns are simulated for 7,125 funds with 36 monthly return observations per fund. Independent monthly returns are drawn from a normal distribution with mean 0.75% and standard deviation 5%. Initial fund assets (TNA_0) equal 10,000 for each fund. Subsequent net assets are calculated endogenously using simulated returns and net cash flows, with $TNA_{jt} = TNA_{j(t-1)}(1 + r_{jt}) + NCF_{jt}$, where r_j is fund j's return in month t, and NCF_{jt} is the end-of-month net cash flow for fund j in month t. In Scenario 1, monthly net cash flows as a percentage of TNA are equal to a random liquidity component with mean zero and standard deviation 1%. Monthly net cash flows in the four behavioral scenarios (Scenarios 2–5) consist of the same random liquidity component plus a behavioral component. The behavioral components are as follows. Scenario 2: symmetric return-chasing behavior in which the behavioral net cash flow component equals 3 times the difference between the fund's lagged monthly return and its mean return; Scenario 3: positive return-chasing behavior only, in which the behavioral component is zero if the fund's lagged monthly return is below the mean and equal to 3 times the difference between the lagged return and the mean return if the difference is positive; Scenario 4: negative return-fleeing behavior, in which the behavioral component is zero if the fund's lagged return is above the mean and equal to 3 times the difference between the lagged return and the mean return if the difference is negative; Scenario 5: contrarian behavior in which the behavioral component is equal to −3 times the difference between the lagged return and the mean return. Returns are percent per month.

ter above average returns, and buy funds after below average returns. Total net assets for each fund are calculated using the simulated monthly returns and net cash flows.[6]

Table 9 reports summary statistics for the average fund return, average investor dollar-weighted return, and the performance gap. In addition, we calculate the dollar-weighted returns on the signed cash flows and report the difference. With random returns and net cash flows (specification 1), the performance gap is zero, as is the difference between positive and negative cash flow returns. The performance gap is positive for the three momentum based strategies in specifications 2–4, indicating poor timing ability, and is negative for the contrarian strategy, indicating positive timing ability. In addition, the difference between returns on positive and negative net cash flows is negative for specifications 2–4, indicating that the average opportunity cost of withdrawn funds exceeds the average return earned on new investments. Results which most closely approximate the actual data are those in Specification 4, where investors flee from low returns, but cash flows to good-

[6] Total net assets at the beginning of the simulation (TNA_0) is set to 10,000 for each fund. Subsequent monthly total net asset values are calculated endogenously using the simulated returns and net cash flows, with $TNA_{jt} = TNA_{j(t-1)}(1 + r_{jt}) + NCF_{jt}$, where r_{jt} is fund j's return in month t, and NCF_{jt} is the end-of-month net cash flow for fund j in month t, specified in (9b), (9c), (9d) and (9e) above.

Table 10. Bond fund and money fund returns

	Mean	Median	25th percentile	75th percentile	Standard deviation
Panel A: Bond funds					
Arithmetic monthly return	0.44	0.44	0.35	0.53	0.29
Geometric monthly return	0.43	0.44	0.34	0.52	0.30
Dollar-weighted monthly return	0.41	0.41	0.31	0.51	0.34
Performance gap	0.02	0.02	−0.02	0.07	0.18
(t-stat.)	(9.44)				
Panel B: Money market funds					
Arithmetic monthly return	0.24	0.25	0.16	0.32	0.13
Geometric monthly return	0.24	0.25	0.16	0.32	0.14
Dollar-weighted monthly return	0.24	0.23	0.13	0.32	0.13
Performance gap	0.004	0.0007	−0.01	0.02	0.05
(t-stat.)	(4.18)				

Bond and Money Market fund datasets include all domestic bond and money market funds in the CRSP Survivor-Bias Free US Mutual Fund Database with returns over the 1991–2004 period. The bond sample consists of 7,222 funds, while the money market sample contains 2,730 funds. For each fund, we calculate the average monthly arithmetic, geometric and dollar-weighted returns over the entire sample period. Performance gap is the difference between fund geometric and dollar-weighted returns. Panel A reports statistics for the sample of bond funds. Panel B reports statistics for money market funds.

performing funds are random. This would also seem consistent with stylized evidence in the literature that poor performance persistence is more easily recognized than superior performance. Overall, the simulation results in Table 9 show that investor return-chasing behavior is broadly consistent with the negative timing ability and performance gap found empirically.

4.3. Timing performance in alternative asset classes

Our analysis so far has focused on the timing ability of equity fund investors, for which we have documented substantial underperformance. However, these results may not necessarily extend to other asset classes having different return and risk characteristics. For comparison with equity fund investors, in this section we examine investor timing ability for two alternative asset classes – bond funds and money market funds.

The bond fund sample consists of all domestic bond funds in the CRSP Survivor-bias Free Mutual Fund database with a unique dollar-weighted average return over the 1991–2004 period.[7] There are 7222 such funds with an average TNA of $228 million, an average load of 2.40%, and average annual expenses of 1.05%. The sample of money market funds includes 2730 funds with unique dollar-weighted average returns over the 1991–2004 period, with an average TNA of $825 million, an average load of 0.26%, and average annual expenses of 0.60%. Performance results are reported in Table 10.

We find that there is a smaller performance gap among bond funds. We also note that the average bond fund returns are much lower than the opportunity cost of withdrawn equity funds. Specifically, the average geometric return among bond funds is 0.43% per month, while the average dollar-weighted return is 0.41%, producing an average monthly

[7] We again note that individual fund share classes are treated as distinct funds for purposes of this study.

performance gap of 0.02%. Moreover, the dollar-weighted return on positive bond cash flows is 0.42% per month, while the return on negative bond cash flows is only slightly higher, at 0.45%, which again indicates only modest mis-timing. Similar results are found for money market funds, where the average performance gap is only 0.002% per month. The average return on positive cash flows is slightly higher than the return on negative cash flows (2.54% *vs.* 2.45% annually).

Overall, the average performance gap is small among bond funds and flat among money market funds, suggesting that negative timing is largely a phenomenon exhibited by equity fund investors. In light of the relatively lower return volatility for bond and money market funds, these results are also consistent with the behavioral story presented earlier.

5. Conclusions

Our study examines the timing ability of mutual fund investors using cash flow data at the individual fund level. We do this by computing the dollar-weighted return earned by investors in each individual fund over the period 1991–2004 and find that the average active fund investor substantially underperforms the growth of a dollar invested in the fund over the entire measurement period. This phenomenon is not only significant for the entire sample but is also found to be robust across various sub-categories of funds whether sorted by size, objective, or risk-adjusted performance. As demonstrated through simulation, this timing underperformance is consistent with investor return-chasing behavior. Furthermore, a comparison of the performance of index fund investors to that of non-index fund investors shows that both groups substantially underperform due to poor timing decisions. This suggests that a significant number of investors who have decided to take a passive approach to security selection by indexing are not necessarily passive in the timing of their cash flows, perhaps preferring a pure timing strategy through this low-cost vehicle.

Certain fund characteristics such as load fees, turnover, and age are directly correlated with an underperforming active investor clientele. It may be the case that more-sophisticated investors are able to locate newer funds to move into as they become available, whereas older and larger funds enjoy significant patronage due to name-brand or a less-mobile or captive investor clientele. Most interesting, however, is the finding that investors who select the best performing funds also exhibit the worst timing performance of all. Thus return-chasing can be a costly endeavor, even when a good fund is found.

Overall, our results suggest that a note of caution is in order for fund investors who are considering whether to attempt market timing. Rather than outperforming a given fund, the average active investor is more likely to underperform a passive dollar invested in the fund. In fact, given the magnitude of average underperformance of new cash flows we have documented, losses from poor market timing decisions likely would erase any potential gains from investing in an otherwise superior fund.

References

Ber and Ruenzi, 2006— Ber, S., and Stefan, R., 2006. On the usability of synthetic measures of mutual fund net-flows. Center for Financial Research Working Paper 06-05.

Bergstresser *et al.*, 2006— Bergstresser, D., Chalmers, J., and Tufano, P., 2006. Assessing the costs and benefits of brokers in the mutual fund industry, Working Paper.

Bollen and Busse, 2001—Bollen, N. P., and Busse, J. A. On the timing ability of mutual fund managers, *Journal of Finance* 56 (2001), pp. 1075–1094.

Braverman *et al.*, 2005— Braverman, O., Kandel, S., and Wohl, A., 2005. The (Bad?) Timing of mutual fund investors, Working Paper.

Carhart, 1997— Carhart, M. On persistence in mutual fund performance, *Journal of Finance* **52** (1997), pp. 57–82.

Chang and Lewellen, 1984— Chang E. C., and Lewellen, W. G. Market timing and mutual fund investment performance, *Journal of Business* **57** (1984), pp. 57–72.

Cumby and Modest, 1987— Cumby, R. E., and Modest, D. M. Testing for market timing ability, *Journal of Financial Economics* **19** (1987), pp. 169–189.

Daniel *et al.*, 1997— Daniel, K., Grinblatt, M., Titman, S., and Wermers, R. Measuring mutual fund performance with characteristic-based benchmarks, *Journal of Finance* **52** (1997), pp. 1035–1058.

Dellva, 2001— Dellva, W. L. Selectivity and market timing performance of fidelity sector mutual funds, *Financial Review* **36** (2001), pp. 39–54.

Elton and Gruber, 1991— Elton, E. J., and Gruber, M. J. Differential information and timing ability, *Journal of Banking and Finance* **15** (1991), pp. 117–131.

Elton *et al.*, 2001— Elton, E. J., Gruber, M. J., and Blake, C. A first look at the accuracy of the CRSP mutual fund database and a comparison of the CRSP and Morningstar mutual fund databases, *Journal of Finance* **56** (2001), pp. 2415–2430.

Fama and French, 1993— Fama, E., and French, K. Common risk factors in the return on bonds and stocks, *Journal of Financial Economics* **33** (1993), pp. 3–53.

Fischoff *et al.*, 1982— Fischoff, B., Slovic, P., and Lichtenstein, S. Calibration of probabilities: The state of the art to 1980. *In*: D. Kahneman, P. Slovic, and A. Tversky, Editors, *Judgement under Uncertainty: Heuristics and Biases*, Cambridge University Press, Cambridge (1982).

Frazzini and Lamont, 2006— Frazzini, A., and Lamont, O., 2006. Dumb money: Mutual fund flows and the cross-section of stock returns. NBER Working Paper w11526.

Graham and Harvey, 1996— Graham, J. R., and Harvey, C. R. Market timing ability and volatility implied in investment newsletters' asset allocation recommendations, *Journal of Financial Economics* **42** (1996), pp. 397–421.

Griffn and Tversky, 1992— Griffn, D., and Tversky, A. The weighing of evidence and the determinants of confidence, *Cognitive Psychology* **24** (1992), pp. 411–435.

Gruber, 1996— Gruber, M. Another puzzle: The growth in actively managed mutual funds, *Journal of Finance* **51** (1996), pp. 783–810.

Henriksson, 1984— Henriksson, R. D. Market timing and mutual fund performance: An empirical investigation, *Journal of Business* **57** (1984), pp. 73–96.

Ippolito, 1992— Ippolito, R. Consumer reaction to measures of poor quality: Evidence from the mutual fund industry, *Journal of Law and Economics* **35** (1992), pp. 45–70.

Kahneman and Tversky, 1972— Kahneman, D., and Tversky, A. Subjective probability: A judgment of representativeness, *Cognitive Psychology* **3** (1972), pp. 430–454.

Kleiman *et al.*, 1996— Kleiman, R. T., Sahu, A. P., and Callaghan, J. H. The risk-adjusted performance of investment advisors: Empirical evidence on selectivity and timing abilities, *Journal of Economics and Finance* **20** (1996), pp. 87–98.

Lee and Rahman, 1990— Lee, C., and Rahman, S. Market timing, selectivity and mutual fund performance: An empirical investigation, *Journal of Business* **63** (1990), pp. 261–278.

Nesbitt, 1995— Nesbitt, S. Buy high, sell low: Timing errors in mutual fund allocations, *Journal of Portfolio Management* **22** (1995), pp. 57–60.

Odean, 1998— Odean, T. Volume, volatility, price and profit when all traders are above average, *Journal of Finance* **53** (1998), pp. 1887–1934.

Sapp and Tiwari, 2004— Sapp, T., and Tiwari, A. Does stock return momentum explain the "smart money" effect?, *Journal of Finance* **59** (2004), pp. 2605–2622.

Shleifer and Vishny, 1997— Shleifer, A., and Vishny, R. The limits of arbitrage, *Journal of Finance* **52** (1997), pp. 35–55.

Volkman, 1999— Volkman, D. A. Market volatility and perverse timing performance of mutual fund managers, *Journal of Financial Research* **22** (1999), pp. 449–470.

Zheng, 1999— Zheng, Lu. Is money smart? A study of mutual fund investors' fund selection ability, *Journal of Finance* **54** (1999), pp. 901–933.

46

Madison Avenue Meets Wall Street: Mutual Fund Families, Competition and Advertising

by

Steven Gallaher
University of Texas
Department of Finance
Austin, TX 78712
(512) 232-2326
FAX: (512) 471-5073
sgallaher@mail.utexas.edu

Ron Kaniel
Duke University
Fuqua School of Business
134 Towerview Drive
Durham, NC 27708
(919) 660-7656
Fax (919) 660-7990
Ron.kaniel@duke.edu

and

Laura Starks
University of Texas
Department of Finance
Austin, TX 78712-1179
(512) 471-5899
FAX: (512) 471-5073
LStarks@mail.utexas.edu

January 2006
The authors would like to thank participants in the Wharton Financial Institutions Center mutual fund conference, the Financial Management Association meetings and seminars at Duke University, Emory University, HEC, INSEAD, the University of Texas at Austin, the University of Texas at San Antonio and the University of Virginia. The authors would also like to thank Keith Brown, Roger Edelen, Stuart Gillan, Jay Hartzell, Woodrow Johnson, Brian Reid, Jonathan Reuter, Stefan Ruenzi, Rob Stambaugh, Ron Wilcox, Hong Yan and Eric Zitzewitz for helpful suggestions on improving the analysis.

Madison Avenue Meets Wall Street: Mutual Fund Families, Competition and Advertising

Abstract

We examine the effects of mutual fund families' strategic decisions, particularly the advertising decision, on investor flows into the families. We find evidence that beyond performance a family's strategic decisions such as advertising, distribution channels, fund offerings and expense ratios, have significant effects on investor flows. Consistent with evidence at the individual fund level, investor flows have an increasing and convex relation to a family's past performance for high performing families. We also find a similar association between a family's flows and its relative levels of advertising expenditures with a significant positive effect for high relative advertisers only.

Madison Avenue Meets Wall Street:
Mutual Fund Families, Competition and Advertising

Investment companies, particularly open-end mutual funds, have been the fastest growing segment of the institutional investor community in recent years. Despite the importance of this institutional investor class, questions still remain concerning the supply and demand for the financial services they provide. Regarding the supply, although early research tended to focus on individual funds (primarily growth funds), more recent research has considered the mutual fund family as the relevant unit of measure.[1] This is an important distinction because many decisions are made from a family, rather than individual fund, perspective since most mutual funds are managed by an investment advisory company that manages a family of such funds. Decisions such as advertising budget, what and when to advertise, the types and number of funds to offer, which distribution channels to pursue, service quality, or individual manager appointments primarily originate on the mutual fund family level. Thus, to fully understand the motivation and impact of these types of decisions, one needs to focus on the mutual fund family complex.

In this paper we address the supply and demand issues by testing hypotheses concerning how the fund family strategic decisions affect investor demand. That is, we examine how the family's strategic decisions affect aggregate net flows of assets under management. Although the investment management company that sponsors the funds is certainly interested in the level of flows to each of their individual funds, they view those funds as a series of products, with the central interest being in the aggregate flows to the entire family of funds. We focus in particular on the advertising decision, but in

[1] Recent papers on mutual fund families include Khorana and Servaes (1999, 2003), Mamaysky and Spiegel (2002), Gaspar, Massa and Matos (2003), Massa (2003), Sigglekow (2003), Nanda, Wang and Zheng (2004), Elton, Gruber and Green (2004) and Kempf and Ruenzi (2004, 2005).

order to compare the determinants of family-level flows to the determinants of individual fund flows found in previous research, we first ignore the advertising decision and examine how aggregate family flows are affected by the outcome of the other family-level strategic decisions. [2]

In considering how to aggregate individual fund relative performance to the family level, we face two considerations. The first is that individual funds compete within their asset class categories for investor funds. The second is that individual funds also compete across asset class categories (even within the same fund family) for investor funds. Thus, the empirical tests can be constructed on two different bases. The first is to simply employ raw returns and the second is to aggregate performance by adjusting returns by the average returns for other funds in the asset class. We employ both analyses and find for both analyses that at the top end of performance, the family flows have a convexity similar to that found in earlier studies on individual growth funds (e.g., Sirri and Tufano, 1998; Chevalier and Elison, 1997). Only the highest performing fund families, overall or relative to other funds in their individual funds' objective classes, have flows related to their performance. When employing raw returns, we find an additional result, family flow and performance are also related for the lowest performing group as well. [3]

Our results suggest that a family's strategic decisions beyond the portfolio management decision can significantly affect investor flows to the family. For example,

---

[2] Previous studies on demand for mutual funds tend to focus on the individual fund rather than the mutual fund family. See, for example, Ippolito (1992), Capon, Fitzsimons, and Prince (1996), Gruber (1996), Goetzmann and Peles (1997), Chevalier and Elison (1997), Sirri and Tufano (1998), Edelen (1999), Barber, Odean and Zheng (2002), and Wilcox (2002). Those studies that have focused on demand for the family of funds use different metrics than we do here. Khorana and Servaes (2003) examine determinants of family market shares rather than fund flows. Thus, they examine percentage levels of assets, while we examine percentage changes in levels of assets. Siggelkow (2003) examines the dollar amount of family flows on an annual basis. We employ the percentage flows, which is comparable to the earlier work on individual funds' flows.

[3] This result contrasts with that of the Sirri and Tufano (1998) individual fund results in that they do not find the relation for the lowest performing funds. For mature funds the results in Chevalier and Elison (1997) are consistent with the ones in Sirri and Tufano (1998). However, for very young funds they also find a positive relationship for poorly performing funds.

the choices mutual fund families make regarding fund offerings has an effect on family flows, consistent with the implications of Khorana and Servaes (1999,2003), Massa (2003), and Gaspar, Massa and Matos (2003). Specifically, our results show that the more objective classes that a families' funds span, the larger the flows to the family, ceteris paribus.

Distribution channels make a difference as well in that flows to the family are significantly related to both load fees and 12b-1 fees. Flows are positively related to the existence of a load, but then decreasing in the size of the load. Flows are also positively related to the family's average 12b-1 fees. The implication of these results is that the choice of distribution channels has significant effects on flows into the family. Families that pay for distribution tend to have higher flows during our sample period.[4]

We find that investors appear to be sensitive to a family's average operating costs. The flow into a family of funds has a significantly negative relation with the family's average expense ratio (excluding 12b-1 fees). This result is at variance with a number of studies of individual funds, which have found no significant relation between flows and expense ratios.[5] Our results suggest that investors pay attention to fees when selecting a family in which to purchase funds.

We then focus on the advertising decision and its effects on individual investors' choice of funds and the aggregate flows of funds into the family. The economic role of advertising in consumer choice problems has been hypothesized to result in the lowering of consumer search costs because advertising provides the consumer with information about the product – such as the product's existence and characteristics (Bagwell and Ramey, 1994). Advertising has also been hypothesized to be a signal of product quality (Nelson, 1970, 1974; Kihlstrom and Riordan, 1984; Milgrom and Roberts, 1986). The

[4] See Bergstresser, Chalmers and Tufano (2004) for an analysis of distribution channels from the shareholders' perspective.
[5] Our result of a negative relation between fund flows and expense ratios is consistent with the findings by the ICI that equity funds with below-average expense ratios hold nearly ninety percent of assets at the end of our sample period (Investment Company Institute, 2005).

implication of these hypotheses is that advertising increases sales of a product due to lowered search costs either for the product in general or for the high quality product. In a similar vein, the attention (or familiarity) hypotheses (e.g., Kahneman, 1973; Merton, 1987; Lee, 1992; Barber and Odean, 2003) imply that advertising has the ability to increase investors' awareness of a fund or fund family, which would be expected to affect mutual fund flows.

Under these hypotheses, the result of a mutual fund family's advertising expenditures would be an increase in assets under management, ceteris paribus, which in turn implies motivation for the use of advertising. It then follows that because the fund family sponsor's income is commonly a percentage of assets under management,[6] increased investor flows from advertising would result in a corresponding increase in the sponsor's income. Additional benefits of the increased investor flows from advertising would include the benefits of economies of scale since the shared expenses for fund management operations would become a smaller fraction of the assets under management, ceteris paribus.[7,8]

Surprisingly little attention has been paid to the link between advertising and fund flows. Since the decision is made at the family complex level, rather than the individual fund level, an examination needs to be conducted regarding how this decision affects family flows. A few studies have examined, at the individual fund level, marketing costs through 12b-1 fees (e.g., Khorana and Servaes, 2003; Barber, Odean and Zheng, 2003) or total fees (e.g., Sirri and Tufano, 1998). Such fees, however, do not reflect the differences in advertising expenditures across mutual funds or mutual fund families. For example, many mutual funds do not charge 12b-1 fees, yet they advertise. Further, Reid

[6] See, for example, Deli (2002), Golec (2003) or Golec and Starks (2004).

[7] Baumol, et. al. (1980), Collins and Mack (1997) and Latzko (2001) have found evidence of economies of scale in the mutual fund industry.

[8] Advertising could also be used as an attempt by fund management to create barriers to entry as suggested by Tirole (1995). Such barriers may be desired since previous research (Baumol, Goldfeld, Gordon and Koehn, 1980; Khorana and Servaes, 2003) suggests that the mutual fund industry is highly competitive with low barriers to entry.

and Rea (2003) cite an Investment Company Institute survey finding that less than five percent of 12b-1 fees were used for advertising and other sales-promotion activities (the remainder was used for distribution charges). Much of the fund family's advertising expenditures are paid by the management company, rather than being a direct expense to fund shareholders through 12b-1 fees. Consequently the full extent of advertising expenditures are not observable through regulatory filings or other common mutual fund databases.

Several studies examine advertising in relation to managed funds. For example, Jain and Wu (2000) provide evidence that the existence of an advertisement in one of two business periodicals is associated with larger flows to the individual funds advertised than to a matched sample of funds.[9] Cronqvist (2005) examines advertising in Sweden by 401-k type funds around the time Sweden launched a partial privatization of their social security system (similar to that being proposed for Social Security in the U.S.). He finds a relation between the funds that advertised and the investors' subsequent allocation choices. Reuter and Zitzewitz (2005) examine advertising expenditures by fund families and conclude that such expenditures may influence mutual fund recommendations of personal finance magazines. Our study differs from these previous and contemporaneous studies in terms of focus, data and methodology.[10]

Fund family complexes typically budget their advertising expenditures and enter into advertising contracts on an annual complex-wide basis. They then often make the decisions regarding when to advertise, and which funds to advertise, later in the fiscal year. Thus, although the advertisement itself may focus on a particular fund, the decision on when to place the ad, how many ads to place, and which funds to include in

[9] Their sample consist of 294 equity funds advertised in Barron's or Money magazines between July 1994 and June 1996.

[10] There has been a surge of recent interest in the relation between operating companies' advertising and their market value or investor interest. See, for example, Frieder and Subrahmanyam (2005), Grullon, Kanatas and Weston (2004), Fehle, Tsyplakov, and Zdorovtsov (2004), and Joshi and Hannssens (2004). Earlier studies in this area include Chauvin and Hirschey (1993).

the ads are made at the family level. It is important to emphasize that the decision is generally not made by individual fund managers. Further, even in the case of ads focused on individual funds, the intent of the ad may be to attract attention to the fund complex rather than simply the fund itself.[11] Because these are complex-wide decisions, it is most appropriate to examine the effects of these decisions on a complex-wide basis and to compare these effects to those of other complex-wide strategic decisions.

With regard to the relation between a mutual fund family's flows and its relative advertising expenditures, we find that advertising affects flows in a non-linear fashion similar to the flow-performance relation, with convexity at the upper end. High relative levels of advertising are significantly related to high fund flows. For low levels of relative advertising there is no significant relationship. Thus, considering the advertising expenditure as a strategic decision, these results imply that for advertising to matter, the family must ensure that they are one of the top advertisers on a relative basis.

Given our evidence regarding the effects of family strategic decisions on flows to the complex, we also examine whether these decisions affect the volatility of those flows. We find that the strategic decisions affect flow volatility in diverse, and significant ways.

An important issue to our analysis and conclusions is the question of the endogeneity of the fund family strategic decisions. Although ultimately all of these decisions are endogenous, because of the lead time required for many of these decisions to have an effect, they are primarily exogenous for our purposes of examining their effects on fund flows. Specifically, the decision on whether to be a high or low fee family is generally not one which would be made overnight as the decision on fees is intertwined with family characteristics such as the quality of services offered or types of

[11] In conversations with mutual fund family executives, they indicated that the intent of the advertising is often not the particular fund advertised, but the fund family itself. Further, investors who call in on the advertisement may be counseled to invest in other funds, depending on their goals and risk tolerances.

funds offered in general (e.g., index or actively managed funds, bond or equity funds). Similarly, the decision on distribution channels is not a fast-changing decision. Thus, in the final section of the paper, we examine the determinants of the family's advertising expenditures. We do so to examine whether there exists a problem with endogeneity and to examine how the other fund family decisions, although exogenous at the point of the advertising decision, may affect the advertising decision itself. Importantly, we do not find that advertising is endogenous to flows, but we do find that the level of the advertising expenditures is dependent on several fund family characteristics: the average expense ratio of the fund family, the distribution channel, the fund's trading costs (as proxied by average turnover) and the fund size.

The remainder of the paper is organized as follow. In the next section we describe our advertising and fund data. In Section II we provide a cross-sectional analysis of the determinants of family flows, followed in Section III by the same analysis with advertising expenditures included as an independent variable. In Section IV we present the evidence regarding whether strategic decisions have an effect on flow volatility at the family level. In Section V we provide results on the determinants of advertising expenditures and we conclude in Section 7.

I. Data

We obtain monthly information on the print advertising expenditures of mutual fund families over the 1992-2001 time period from Competitive Media Research (CMR). CMR is a third-party collector and distributor of data on advertising expenditures for many products, both print advertising and other media advertising.[12] Over our sample period these advertisements appeared in over 288 publications, from the *Wall Street Journal* (the greatest amount of advertising dollars spent) to the *Elgin Courier News* (the

[12] Our data is limited to the print advertising, but according to Reuter and Zitzewitz (2004), mutual fund print advertising accounts for about 80% of total advertising expenditures.

least amount of advertising dollars spent). Our data on characteristics of mutual fund families and their constituent funds (such as total net assets, expense ratios, load fees, 12b-1 fees, objectives and returns) is obtained from the CRSP mutual fund database. Since our focus is on the mutual fund family rather than individual funds, we only include families with a minimum of $1 billion under management. Although there are a number of very small mutual fund families (see, for example, Kempf and Ruenzi, 2005), this restriction has little impact on our sample as the fund families with at least $1 billion in assets under management covers 99.5% of the total net assets of mutual funds that advertised in the CMR database at the end of our sample period (2001) and 97% at the beginning (1992).[13]

Table 1 provides characteristics of the mutual fund families included in our sample over the 1992-2001 time period. Consistent with the changes in mutual fund assets in general over the sample period (see Figure 1), the number of large fund families grows from 98 in 1992 to 124 in 1996 and then contracts to 109 in 2001. Figure 1 and Table 1 indicate that the total assets under management at these families increases from $935 billion to over $5 trillion, ending at about $4.2 trillion. Consistent with the mergers of mutual fund complexes in the late 1990s, there appears to have been some consolidation in the industry. However, 109 mutual fund families with over $1 billion in assets under management are still remaining at the end of our sample period. Thus, it is perhaps not surprising that an analyst for the mutual fund industry stated that the "degree of fragmentation is greater today than it was in 1990, contrary to other parts of the financial services industry."[14] Consistent with this statement, Table 1

[13] We omit the very small fund families because their differences from the typical fund family (including the small assets under management, the small number of funds offered, and the lack of capability for advertising) potentially results in a decision process that would vary considerably from that of the large fund families. The fact that very few of the small fund families show up in our advertising database supports our assumption that they lack the capability of advertising in the same manner as the larger fund families.
[14] Presentation by Guy Moszkowski at Wharton Financial Institutions Center's "Mutual Fund Portfolios in Theory and Practice" Conference, May 7, 2004.

shows that although the average assets under management for a fund family grew from $9.54 billion in 1992 to $38.76 billion in 2001, a four-fold increase, the share that this represented of the total market fell from 1% to 0.9%. The table also shows the growth in assets under management was strong in the early 1990's, with average monthly net flows of about 4% of assets, but these flows fell to less than 2% in 2001.

Table 1 provides information on distribution channels of the families in the sample. The broad use of 12b-1 fees and load fees implies that mutual fund complexes, in general, use multiple distribution strategies. At the beginning of the sample period, almost 70% of the families had at least one share class that charged 12b-1 fees. By the end of the sample, 83% of the families had at least one share class with 12b-1 fees. Similarly, at the beginning of the sample period about 70% of the families had at least one fund with a front-end load fee, by the end of the sample period, 75% of the families charged such fees. Thus, not only do most families use multiple channels, but the use of multiple channels has been increasing over time. Across all funds in a family, the average front-end load fee was about 1.82% in 1992 and 1.61% in 2001, however, when we restrict the average to funds within a family that have a load, there is little change in the average load fee across time, remaining between 4% and 5%. The difference is due to the offering of more funds without load fees.

Table 1 shows a small increase in average expense ratios over the ten-year period. This increase in average expense ratios for the fund families is most likely due to an increase in specialized or international funds over the period, which have higher costs of operations.

II. Determinants of fund flows on the family level

In this section we provide a cross-sectional analysis of the determinants of fund flows on the family level using the same explanatory variables as in previous research on the individual fund level. This allows us to compare the results for family flows to those of previous studies that employ individual funds.[15]

For each month, the dependent variable is the net flows into fund family *k* for month *t*:

$$NetFlow_{k,t} = \Sigma_i \{TNA_{i,t} - (TNA_{i,t-1} * (1+R_{i,t}))\}/TNA_{k,t}$$

where $TNA_{i,t}$ represents fund i's total net assets at time t, R_{it} represents fund i's return in month t and $TNA_{k,t}$ represents fund family k's total net assets at time t. Figure 2 shows more detail regarding net fund flows over the sample period. As the figure shows, quite a bit of volatility exists in family flows over the sample period. Such large volatility suggests that a mutual fund complex may be able to influence the flows into their complex through their strategic decisions.

A. Aggregating raw returns

The primary independent variable employed in previous research on determinants of flows into individual funds is the past return performance of the fund. Accordingly, our primary independent variable in this section is the fund family's average return performance over the previous year. A complicating factor is that he individual funds within a family compete both in general and in different segments of the market with different asset classes and returns. Consequently, we employ two different methods for aggregating individual fund performance to the family level. In this subsection, we measure family return performance as the average return on the individual fund portfolios, weighted by the total net assets (i.e., market value) of the funds. In subsection IIB, we measure family by aggregating each individual fund's excess return over their objective category average.

---

[15] Examples include Ippolito (1992), Gruber (1996), Goetzman and Peles (1997), Chevalier and Ellison (1997), Sirri and Tufano (1998), Lynch and Musto (2003) and Huang, Wei and Yan (2004). A reverse relation (the effects of fund flows on future returns) has also been considered, theoretically by Berk and Green (2003) and empirically by Wermers (2003).

Because we need to aggregate the monthly cross-sectional regression results across time, we normalize the family average returns within each period. To do so, we follow the Sirri and Tufano (1998) technique of ranking the sample average returns over the immediate past year and then normalizing these rankings onto the [0,1] interval.[16] The advantage of this technique is that it converts the family's average returns into their rankings in comparison to other families' returns on a period-by-period basis.

Our strategic decision variables for this regression are: the log of the number of fund objectives offered by the family (the maximum is 17), a dummy variable to indicate fund families that have at least one fund with a load fee, the average front-end load fee ranked against other families, a dummy variable to indicate fund families that have no 12b-1 fees, the average 12b-1 fee ranked against other families, the average expense ratio (excluding 12b-1 fees) ranked against other families, and the average turnover of the funds' portfolios as a proxy for fund trading costs.[17,18]

We also include several control variables in the regression. Since a potential complicating factor in the empirical specification of the model is the existence of persistence in fund flows, we employ lag fund flows (i.e., the fund's flows over the previous month). Given prior evidence showing that star performance results in greater cash inflow to the fund and to other funds in the family, we include a dummy variable equal to one if the family has a star fund in the month (where star fund is defined as a fund whose return is in the top five percent of returns for the fund's category for the past

---

[16] That is, each family is assigned a number between 0 and 1, with the best performer getting a 1 and the worst a 0. In between, the numbers are evenly spaced.

[17] We average return, expense ratio, 12b-1 fees, load fees, and turnover by calculating the market-weighted average across funds in the family.

[18] Another strategic decision is whether to waive part of the fund's fees (see Christoffersen, 2001). As we do not have data on fee waivers, we do not explore this decision from a family level perspective. It should be noted that since CRSP reports actual expense ratios, such decisions are imbedded in our results.

year).[19] We also control for family size through the log of the total net assets (at the beginning of the month).

No theory exists to give us guidance as to the correct specification for the fund flow-performance relation. Previous empirical studies at the individual fund level have employed a variety of specifications, with a large number of studies providing evidence of a nonlinear relation (e.g., Ippolito, 1992; Carhart, 1994; Gruber, 1996; Chevalier and Elison, 1997; Goetzmann and Peles, 1997; Sirri and Tufano, 1998; and Lynch and Musto, 2003).[20] Thus, we employ a specification that allows for this nonlinear relation. Specifically, we employ the Sirri and Tufano (1998) piecewise linear specification using cross-sectional regressions on a monthly basis and assuming that the kinks are identical across the months. Once we have run the cross-sectional regressions for each month, we then use the Fama-MacBeth (1973) technique to aggregate the coefficients across the 1992-2001 sample period.[21] The Fama-MacBeth method is the most appropriate approach for our pooled times series data because it allows for differences in relative advertising across periods and controls for seasonality in advertising. Using a panel data set with fixed effects would not allow the slopes to change over time.

In Table 2, we show the results of two piecewise linear specifications, one of which (Model 1) assumes four kinks in the specification and the second of which (Model 2) assumes two kinks.[22] The results from these specifications show a strong relation between the flows into a mutual fund family and the family's past average return performance if the family exhibits extreme return performance. In Models 1 and 2

[19] Nanda, Wang and Zheng (2002) define star fund as a fund in the top 100 performers of a category. They further state that the star funds constitute about 5% of their sample. Such funds should also be related to funds with top Morningstar rankings as Morningstar ratings are heavily dependent on returns (Blume, 1998; Sharpe, 1998; Del Guercio and Tkac, 2002).
[20] Most previous studies have examined individual fund flows on an annual basis (e.g., Sirri and Tufano, 1998), semiannual basis (Edelen, 1999), or aggregate flows on a daily basis (e.g., Edelen and Warner, 2001). Previous studies have not examined the determinants of flows into a family of funds on a monthly basis as we do here.
[21] All of the Fama-MacBeth (1973) t-statistics are based on the Newey-West (1987) heteroskedasticity and autocorrelation consistent standard errors.
[22] Allowing for one kink or three kinks does not change our conclusions.

monthly family flows are significantly related to whether the family's average return in the previous year is in the bottom or top group of all families' average returns. If a mutual fund family is in the one of these extreme performance groups, the models suggest that the flows will be positively related to that performance.

In general, previous studies find that the flows to individual funds are related to the fund's past performance, but more so for the highest-performing funds than the lowest-performing funds. In contrast to our analysis of the value-weighted family flows, most of these studies have restricted their samples to individual funds with a growth objective. The relation we find at the upper end of the return distribution is consistent with the previous results for individual funds (e.g., Sirri and Tufano, 1998; Chevalier and Ellison, 1997). The relation at the lower end of the return distribution is consistent with the earlier Chevalier and Ellison results, but not with the earlier Sirri and Tufano results. Further, the magnitudes of the coefficients and t-statistics on the top and bottom performance groups suggest a stronger relation for the top performers than the bottom performers. The positive coefficient for the worst fifth (Model 1) or worst third (Model 2) of performers suggests that investors do respond (and leave), the worst performing fund families.

Some fund families specialize in certain categories of mutual funds such as fixed income funds. Even without such specialization, the proportions (and net assets) in the different fund categories vary across families, suggesting that our results on the flow-performance relation at the family level could be driven by the different proportions of fund categories in the families. This concern is supported by Lettau's (1997) analysis in which he correlates mutual fund flows with lag returns for different categories of funds and finds stronger correlations for aggressive growth and growth funds than for growth and income or balanced funds. To test whether our family results are influenced by different proportions of equity and fixed income funds across families, we also ran the cross-sectional regressions in Table 2 including only the growth funds in the families.

We found the same qualitative results as we did when including all of the funds. The significantly positive coefficients were again in the bottom and top groups. We investigate this issue further in subsection IIB.

The coefficients on the strategic decisions suggest that these decisions can indeed have an effect on family demand. Consistent with the implications of the family analyses of Mamaysky and Siegel (2002), Khorana and Servaes (1999,2003), Massa (2003) and Gaspar, Massa and Matos (2003), we find that the family fund flows are increasing in the number of objectives in which the family offers funds. One-stop shopping seems to pay off. For example, Mamaysky and Spiegel (2002) develop a model of mutual funds in which the fund families do not specialize, rather the optimal strategy is for the families to offer their products in multiple fund categories. Their model is consistent with empirical and theoretical work by Sigglekow (2003), Massa (2003), and Khorana and Servaes (2003). Sigglekow finds that fund families with more diversified offerings (i.e., less focus) have greater dollar flows. Massa similarly argues that a family's tendency to offer multiple funds across fund categories is a tool that fund families can employ to limit competition and increase market coverage. Khorana and Servaes find that product differentiation is an important aspect of competition among mutual fund families.

Table 2 also provides implications regarding strategic decisions for the family's distribution channels. Being a fund family with at least one fund with front-end load fees is associated with higher fund flows. This result suggests that brokers and financial advisers, who receive the load fees, can increase flows into funds for which they receive a commission. However, the coefficient on the family's ranked front-end load fee shows that fund flows are decreasing in the magnitude of the load fee, implying that larger loads impose an impediment to increased flows. This negative relation is consistent with previous evidence on the deterrents of load fees to mutual fund purchases. For example, Barber, Odean and Zheng (2003) find a negative relation between individual fund flows and the magnitude of the load fees. Thus, although the load fee may be considered a marketing expense to increase flows into the fund, its magnitude has a detrimental effect on flows in the cross-section. That is, the benefit to the fund family comes from having a relation with a broker, but conditional on having such a relation, the higher magnitude load fee has a detrimental effect.

The results for the 12b-1 fees are somewhat different. While there is no difference in flows between families with 12b-1 fees and those families that do not pay such fees, those families that pay a higher magnitude of these marketing fees receive higher inflows. These results on strategic decisions regarding distribution channels indicate that fund families with multiple distribution channels, but low load fees, do the best in terms of increasing their overall fund flows.

We have two proxy variables to capture the operating costs of the mutual funds that are borne by the shareholders: the average expense ratio (without 12b-1 fees) and the average turnover (which should be correlated with the funds' average trading costs). In contrast to much of the research on individual funds, we find that the coefficient on the ranked average expense ratio is significant and negative, implying that investors are sensitive to this source of fund costs on a family level. The coefficient on the other proxy for fund costs, portfolio turnover, has a significantly negative relation with fund flows, indicating that families with higher turnover are less attractive to investors, ceteris paribus.

In terms of the control variables, the coefficients remain at approximately the same sign, magnitude and significance across the two models. The coefficient on the lag flow variable shows a strong persistence in flows to a family across periods. The large persistence in fund flows (about 30% of the previous month's flows) suggests that funds receive a sizable proportion of their flows from fixed commitments such as retirement accounts or savings plans. In addition, larger families receive smaller percentage flows, on average. Consistent with the empirical evidence of Nanda, Wang and Zheng (2002), we find that fund families with a fund in the top five percentile of performance of the funds in their investment category (a star fund) receive a higher inflow of performance, This result is also consistent with the evidence in Del Guercio and Tkac (2002) regarding the effects of Morningstar ratings.

In summary, our specifications for the determinants of fund flows on the family level support the hypothesis that the convexity in the flow-performance relation found for the high-performance individual funds continues for the fund family in aggregate and that there exists a further non-linearity with the lowest performing mutual fund families. Our results also indicate that a family's strategic decisions can have significant effects on investor flows to the family.[23]

B. Aggregating returns adjusted for average objective category performance

In order to consider the competition between funds in the same asset class, in this subsection we examine the flow-performance relation when family performance is aggregated on an adjusted return basis. We adjust individual fund returns for their average objective category performance in the following way. For each individual fund's return, each month the total net asset weighted average return for the associated ICDI objective code is calculated and deducted to obtain the fund's return excess to their market segment average. Then we calculate the average excess return for the family by taking the average excess return on the individual fund portfolios, weighted by the total net assets (i.e., market value) of the funds. Using this measure of families' returns, we then again conduct the analysis in Table 2. The results, presented in Table 3, are similar to Table 2, with a few notable exceptions. There is no longer a significant relation between family flows and performance for the lowest performing group, although the

[23] For an additional robustness check, we also ran cross-sectional regressions in which we used average load fees, 12b-1 fees, and expense ratios, rather than using their ranked values as we have in Table 2. We found no qualitative difference in results in terms of magnitudes or significance.

convexity at the high end still exists. Several of the other variables are no longer significant, log fund objectives offered, star fund in family and family size.

As it is not clear which is the most appropriate specification, in the remainder of the analysis, we employ the raw return specification, noting any differences from the adjusted return specification where appropriate.

III. Advertising and mutual fund flows

As pointed out earlier, the economic role of advertising in consumer choice problems has been hypothesized to result from the lowering of search costs for the consumer. For example, according to Bagwell and Ramey (1994), optimal consumer search uses advertising to guide that search. With regard to mutual funds in particular, Lettau (1997) points out that a rationale for individuals to invest in mutual funds "can be viewed as a decision in terms of an optimal allocation of time." If the investors want to reduce time spent on researching financial markets, they most likely would also want to reduce time spent on researching fund management and performance, particularly given that more mutual fund share classes exist than do stocks traded on the NYSE, AMEX and NASDAQ national markets combined. In fact, Sirri and Tufano (1998) and Huang, Wei and Yan (2004) argue that fund flow should be related to the mutual fund investors' search costs.[24] Similarly, the attention and familiarity hypotheses imply that advertising should increase investor awareness and subsequent fund flows.

A. Advertising and industry flows

Figure 3 presents the dollar amounts spent on mutual fund advertising across our sample period. As the figure shows, there exists quite a bit of variation throughout the time period, with an apparent seasonality. Figure 4 presents the average percentage of each year's advertising expenditures that occurs in a given month, showing that most

[24] Tkac (2004) provides a discussion of mutual fund investor search costs as well.

monthly advertising is relatively higher in the beginning of the year.[25] Before testing the hypothesis that advertising affects family flows, we first test the hypothesis that flows in the mutual fund industry in general are affected by aggregate levels of advertising in the industry. To do so, we regress percentage monthly flows to the industry against the advertising expenditures as a percent of total assets under management (lagged by one month), the total industry flows from the previous month, and the average industry return from the previous year. The results, provided in Model 1 of Table 4, show that there exists a strong positive relation between aggregate industry flows and advertising expenditures. The table also shows a strong positive relation between flows to the industry and returns in the previous period. This later relation is consistent with that found by Edelen and Warner (2001) for daily aggregate flows to equity mutual funds.

Advertising by a mutual fund family has the potential to affect not only its own flows, but flows to other fund families as well. In Model 2 of Table 4, we test this related hypothesis by limiting the dependent variable to flows into families that did not advertise. We find that their flows are affected by advertising by others in the industry as well, suggesting substantial spillover effects from advertising.

4.2 Family fund flows and advertising expenditures

Several previous studies have examined the relation between individual fund flows and proxies for advertising of those funds. For example, Sirri and Tufano (1998) use total fees charged as a proxy for marketing and distribution expenditures.[26] They find no relation between the flow-performance relation and this proxy, except in the case in which they separate the funds into those with high fees and those with low fees. In that case they find that funds with higher fees, which the authors assume are funds with greater marketing efforts, have greater flow-performance sensitivity. However, because

---

[25] Using the Fama-MacBeth (1973) regressions for our analysis controls for this seasonality.
[26] Sirri and Tufano (1998) include one-seventh of any load fee in their total fees charged measure.

they are forced to employ a coarse proxy for marketing efforts, they cannot ensure that their results are not caused by confounding factors, such as funds with higher service levels (associated with the higher fees) attracting greater flows.

Jain and Wu (2002) use a dummy variable approach to compare fund flows of individual funds that have advertisements in one of two magazines in a month to flows of funds without advertisements in these magazines. Over their July 1994 through June 1996 sample period, they find that the advertised funds have higher net inflows, after controlling for prior performance, lag flows, and size. In a study concurrent to ours, Cronqvist (2005) examines a number of issues with the advertising of the Swedish 401k-type funds, including what funds advertise, which types of advertising affect the investors allocation choices, and whether fund advertising is a signal of future performance.

These studies have focused on the role of advertising in individual funds, but it is important to keep in mind that the advertising expenditure decision is a fund complex decision, not an individual fund decision. Thus, we test the hypothesis of whether advertising affects flows on the complex level. Because of the differences in size across the fund complexes (and consequent differences in ability to spend advertising dollars), we need to scale the monthly advertising expenditures. Our choice is to use the total net assets under management for the fund complex. As in our measure of the return variable discussed in the previous section, we need to aggregate the cross-sectional relation between family fund flows and advertising expenditures across the multiple monthly periods. Accordingly, we normalize the advertising variables on a [0,1] interval analogous to the Sirri and Tufano (1998) normalization procedure for the performance variable. We then assume a piecewise linear relation between family flows and advertising expenditures, an assumption similar to the assumption regarding the relation between family flow and past performance. In addition, we include a dummy variable if the family did not advertise during the month.

Our specifications for the flow-performance relation in these analyses are the same piecewise linear specifications employed in the previous section. We also include the same strategic decision and control variables: the log number of objectives that the family's funds span, a load dummy, the ranked average load fee, a dummy for 12b-1 fees, the ranked average 12b-1 fees, the ranked average expense ratio, the average turnover of the funds' portfolios, the lag family fund flow, the log of the total net assets from the previous month, and a dummy variable equal to one if the family has a star fund.[27]

Table 5 shows the results from this analysis. For easier comparisons of coefficients, Model 1 shows the two-kink piecewise linear flow-performance relation without advertising variables from Table 2. Models 2 and 3 include the advertising variables. Model 2 has the simplest linear specification of advertising in which we have a variable for no advertising and a variable for the advertising expenditures ranked against other families. Model 3 employs the piecewise linear specification as described above. The results of Models 2 and 3 show that advertising has a significantly positive effect on fund flows for the heavy advertisers, but that advertising viewed from a simple linear specification appears to have no significant effect.[28] These results suggest that a threshold of advertising expenditures relative to competitors' advertising expenditures exists before the advertising has significant effects on flows into the family. The advertisers who spend the least have no significant relation between their advertising dollars and fund flows. In contrast, those families in the middle range of advertising spending per assets under management, show either a negative or no flow-advertising relation. Given that advertising has a significantly positive impact only at the top end, the advertising decision becomes a strategic decision for the fund family management.

---

[27] In these cross-sectional analyses we omit any fund families that are less than three years in age or that have only three months of advertising expenditures over the entire sample period.
[28] In a separate analysis (not reported) we run a pooled cross-sectional analysis of the effects of advertising on the fund flows. The results are consistent with those reported in Table 4.

The results imply that just advertising is not sufficient for significantly increasing flows, rather the family has to extensively advertise relative to other families' decisions in that period.

As in the analysis without advertising expenditures in Model 1, the convexity in the flow-performance relation appears for the top performing funds. The strategic decision and control variables still have effects similar to those when the advertising variable is absent. Conceptually, one might expect advertising to affect the flow-performance relation in that advertising could mitigate or magnify the importance of fund performance. We do not find this to be the case. Comparing the coefficients on the performance variables between Model 1 and Models 2 or 3 shows virtually no change when advertising is included in the regressions.[29] Thus, while Table 5 shows that family advertising expenditures can affect family flows, it does so independent of the family return performance. Similarly, the relation between fund flows and advertising also does not affect the relation between flows and the magnitude of the 12b-1 fees. Families with larger 12b-1 fees have higher net inflows regardless of the extent to which they advertise.

The results from Table 5 suggest that mutual fund families can affect net flows through the performance of their funds, including achieving star status for at least one fund in the complex. They also have additional independent strategies with which they can affect their net flows: spend a sufficient percentage of assets on advertising relative to their competitors, offer funds in a large range of objectives, pay marketing expenses for distribution channels through load or 12b-1 fees, or lower their expense ratios.[30]

Our evidence on the effects of advertising and its role would be consistent with the arguments of Massa (2003). In discussing fund family decisions, Massa argues that

---

[29] Adding an interaction term between performance and advertising does not change these conclusions.

[30] We also ran the regression by normalizing the ad variable by the beginning of the year TNA instead of the beginning of the month TNA. We found no change in results.

performance-maximization is not necessarily the optimal strategy for fund families – that the profit-maximizing mix of fees, performance and number of funds could result in lower levels of performance. This results from the ability of fund families to differentiate themselves in terms of non-performance related characteristics so that they do not need to compete solely on the basis of performance.

The levels of significance in Table 5 show that heavy advertisers can increase flows into the family. The economic significance of these results is reflected in Figure 5, which shows the total returns to advertising in terms of the flow/advertising relation. The figure shows the times series results for the 85th and 95th percentile advertisers at each point in time. That is, at each point in time we took the advertising expenditures of the fund family that was closest to (and above) the 85th (95th) percentile in advertising expenditures and applied the coefficients from that period's regression to generate the dollar flows to advertising. The dollar flows to advertising were then divided by the actual advertising cost to derive the returns to advertising. As Figure 5 indicates, the returns to advertising for the heavy advertiser can be economically large.

Thus far, we have found that heavy advertising by a mutual fund family results in statistically and economically significant increased flows to the family. The question that naturally arises is the degree to which the advertising has persistent effects. That is, does the advertising affect individuals who are making investment choices soon or is there a residual effect on individuals who make their choices later? We reran the regressions in Table 5 and included the advertising lagged by one month, or alternatively the cumulative advertising of the previous two months, four months, or twenty-four months. This was done for a linear specification, as well as a piece wise linear specification, for the advertising lagged variables. The results (not shown) indicate that the advertising from previous months has no effect on the flows, suggesting that there is no persistence in advertising – the level of the most recent month's advertising expenditures dominates. That is, advertising is short-lived.[31]

In order to determine whether our results were driven by flows to the star funds in the family (which are more likely to be the advertised funds as well), we reran the regressions in Table 5 eliminating star funds from both the left and righthandside variables. That is, the dependent variable, flows to the family, does not include flows to any star funds, and each of the independent variables is calculated without the inclusion of the star funds. For the advertising variables, the results are qualitatively identical to the results when star funds are included. Heavy advertising relative to other families results in significantly increased flows to non-star funds in the family.

Given the volatility in advertising expenditures across time, we examined whether advertising expenditures have differential effects in up-markets as compared to down-markets. Accordingly, we divide the sample period into those months in which market returns were positive over the month and those periods in which market returns were negative. We find that advertising does not have significant effects during down-markets. Our results are driven by the relation between fund flows and advertising expenditures during up-market periods.[32]

Two potential problems in our analysis could develop from our methodology of scaling the advertising expenditures. The first problem is that our results could be driven by a spurious correlation between fund returns and advertising. That is, since we scale our advertising variable by the previous month's total net assets under management of the fund family, we could be inducing a result between the change in the total net assets and the fund flows. To check whether this potential problem, we reran the regressions in

---

[31] These results are consistent with previous studies for operating companies which have found that advertising has a short-term elasticity on sales that is positive but low (see, e.g., Leone and Schultz, 1980 or others cited in Joshi and Hanssens, 2004).
[32] Over our sample period, only 15% of the months were down months according to this definition. We also define up-markets and down-markets for the market return compared to the riskfree rate. The results remain unchanged.

Table 5 and scaled the advertising expenditures by the total net assets from the beginning of the calendar year. That is, for each month in a year, a family's advertising expenditures is scaled by the same variable, which does not change during the year. Our results from this analysis do not qualitatively differ from those reported, suggesting that our results are not driven by variation in performance rather than advertising expenditures.

The second potential problem is that our scaling methodology is inappropriate because of the differences in sizes across the fund families. To check this problem, we divided our advertisers into three groups, by size of assets under management, and then ranked the advertising expenditures within each group. Again our results are qualitatively the same as those reported.

5. Family fund flow volatility and strategic decisions

Chordia (1996), Edelen (1999), Greene and Hodges (2002), and Rakowski (2003) have suggested that fund flow volatility is costly to mutual fund operations. That is, the uncertainty with regard to the level of the investors' investment in the portfolio can affect the performance of the portfolio. Thus, an important factor in a fund complex's strategic decisions could be the concomitant effects on the complex's average flow volatility across its funds. A priori, the direction of these effects is unclear. On the one hand, the strategic decisions could bring in a constant stream of dollars or result in lower overall redemptions by shareholders in the family complex of funds, thus, reducing flow volatility, ceteris paribus. (For example, Goetzmann and Peles (1997) hypothesize that advertising could discourage shareholder redemptions by reducing their cognitive dissonance.) The mutual fund family could then make strategic decisions in part to manage the cost of their flow volatility. On the other hand, if these decisions successfully attract additional flows to funds, they could also have the unintended consequence of increasing flow volatility and costs. For example, in the case of advertising, the decision could increase flow volatility by attracting additional assets in an uneven fashion, particularly if the advertising is sporadic or targeted toward particular funds based on their previous performance. This could be the case given previous research. Kempf and Ruenzi (2004b) find that a fund's growth is dependent not only on the fund's return relative to its peers, but also relative to other funds in the same family. Such a result would be consistent with families advertising their best funds and those funds having higher growth due to the advertising.[33]

In this section we examine whether the family's strategic choices affect the family's average fund flow volatility in either of these directions. The dependent variable for our tests is the average standard deviation of fund flows over the previous twelve months, where the average is taken across the funds within the family.[34] Given that it is likely the family would be most concerned about the flow volatility in the smaller funds in the family , we employ an equal-weighted average of the individual funds' flow volatility.[35] We control for the persistence in family flow volatility by including the previous year's flow volatility measure. One aspect of family flow volatility is that larger families with more funds being offered could have lower average flow volatility simply because the averaging process could make outliers less important. We control for this effect with two variables: current family flows and the log total net assets. Whether having a star fund in the family adds to the family's flow volatility is an empirical question we address in this analysis by including a dummy variable for whether the family had a star fund in the previous period.

The results of these analyses, provided in Table 6, show that the family's strategic decisions have mixed effects on the family flow volatility. Not advertising

[33] Consistent with this result, Jain and Wu (2002) find that advertised funds previously earned higher returns than their category benchmarks.
[34] The results are similar if volatility is computed over six months rather than twelve months.
[35] For robustness we also looked at the case where fund volatilities are value weighted. The results for the strategic decision variables, including the advertising variables, are similar to the ones reported.

increases flow volatility. However, the relative level of advertising does not seem to significantly impact flow volatility. We do find that load fees, 12b-1 fees, and expense ratios significantly affect the volatility. The existence of a load fee increases flow volatility as does the existence of 12b-1 fees. These results combined suggest that the use of distribution channels increases the variability of the flows into the funds, adding more uncertainty. The size of the load fees reduces the volatility of the flows. Families with higher average load fees tend to have smaller flow volatility as well, suggesting that employing brokers can reduce the volatility of fund flows, thus, reducing the costs to the existing fund shareholders (who do not encounter the front-end load fees). In addition, the size of the family's average expense ratio relative to other fund families reduces the volatility of the flows. The latter could occur if the expense ratios are a proxy for service, as suggested by Sirri and Tufano (1998), and investors are more likely to stay in a fund when there is increased service.[36] The result could also occur if investors who choose the families with greater expenses are more stable investors, either because they value some aspect of the more expensive fund families (e.g., service) or because they are passive investors as suggested by Christoffersen and Musto (2002).

One issue with the fund flow volatility is that if fund families advertise to reduce fund flow volatility, they would be most concerned about reducing the volatility on the downside rather than the upside. To test whether there are asymmetric effects on volatility from advertising, we reran the regressions in Table 6, using semi-variance rather than variance as our dependent variable. The results are shown in Table 7. The results for the strategic decision variables, including the advertising variables, are similar to the ones in Table 6.

[36] The willingness of investors to pay for mutual fund service (or for financial advisers' service) may explain the willingness of such investors to pay differential fees for S&P 500 index funds. and explain the puzzle of the Elton, Gruber, and Busse (2004) results.

VI. Determinants of advertising

One issue that arises from the models of the relation between advertising and family flow is the issue of whether an endogeneity exists in the relation. For example, fund management companies with higher flows, and thus, higher resultant management fees, could have more resources with which they could pay for advertising. We investigate this issue by examining whether systematic determinants exist for a family's choice of the amount of advertising dollars to spend.

Economic studies of firm advertising have proposed various motivations and roles for the advertising.[37] For example, as discussed earlier, advertising can reduce search costs for the consumer. Additionally, Nelson (1970, 1974) argues that for some types of products ("experience" products) the quality of the product is not ascertainable prior to purchase. For such products, the existence of advertising itself can reflect a high-quality product. Nelson further argues that the key to advertising is in repeat purchases for a product. Since consumers are more likely to repeat a purchase of a high-quality product, it becomes important for the high-quality producers to advertise and reach the consumer on the first purchase. In terms of the mutual fund market, the repeat purchases manifest themselves in maintaining and increasing investment in the fund.[38] The quality of a mutual fund family could depend on several factors, including performance and services.

To test determinants of family advertising expenditures, our dependent variable is the relative level of advertising, i.e., the annual advertising dollars spent by the family normalized by the family's total net assets under management. The denominator in the

---

[37] Additional economic models of advertising examine the role of advertising across different industry structures (e.g., Greer, 1971). For analysis of these types of models in thrift and banks, see DeYoung and Ors (2005) and Ors (2005). Because we have data on only one market, which is national, these industrial organization models are outside the scope of this paper.

[38] Nelson's (1970, 1974) hypothesis considers repeat purchases to be the key goal of advertising, however, with regard to mutual funds, Johnson (2004) finds that most individual investors do not make repeat purchases of the same mutual fund. The question of who the advertisements are reaching to increase fund flows is the subject of ongoing research.

dependent variable is lagged by one year because the relation between performance and the current level of assets under management could mask a relation between the advertising variable and performance. The potential explanatory variables are proxies for family quality plus other strategic decision and control variables used in the earlier analyses.

A fund family's quality can be reflected in their return performance and in their investor service. The two measures of return performance that we employ are the current month's return and the previous year's return. Unfortunately a good proxy for the quality of a fund family's service is not available. Sirri and Tufano who posit that total expenses may be a measure of services provided by the fund. Consequently, we employ the fund family's ranked average expense ratio (excluding 12b-1 fees). Our proxies for the distribution channel decisions are a dummy variable for the existence of load fees, the ranked average load, a dummy variable for the existence of 12b-1 fees and ranked average 12b-1 fee.

As we have hypothesized that the other strategic decisions are exogenous to the annual advertising decision (due to their long lead times), we examine whether the proxies for these decisions can help explain the advertising decision. In the long run, the strategic decisions would be potential substitutes for each other.

The control variables in this regression are the previous year's flow, the previous year's flow volatility, the logarithm of total net assets, and ranked average turnover (as a proxy for trading costs). Because economies of scale can affect the ability to advertise as well as the benefits from the advertising, we also divide the sample of families at the median for the size of total net assets under management. We run the cross-sectional regressions on an annual basis and use the Fama-MacBeth (1973) technique to aggregate the coefficients across the periods.[39]

[39] Because of the limited power of the Fama-MacBeth (1973) technique with the annual regressions, we also ran a pooled, cross-sectional regression. There was no increase in the significance of the independent variables – the results were basically the same.

The results when all mutual fund families are included in the regression are shown in Model 1 of Table 8. In Models 2 and 3 the families are divided by the size of assets under management, where assets under management are measured one year prior to the observation.[40] The evidence in Table 8 suggests that endogeneity is not a problem for our earlier results on the flow-advertising relation. We find that neither the family's previous annual flow nor the volatility of the flows appear to influence the advertising budget.

According to all three models, the amount of advertising dollars spent by a family is not affected if the family's relative average current return or return in the previous year (as ranked against other families) is in the highest performers. Thus, it is not the case that when a complex performs well, they advertise more. The one group in which return performance matters is for the small families who are the poorest performers. These families have a tendency to have a significantly higher level of relative advertising.

Across both large and small mutual fund families, the amount of advertising dollars spent by a family is positively influenced by the family's average expense ratio.

Interpreting the results for mutual fund family performance and family expense ratios as reflections of quality would provide conflicting interpretations. If one assumes that quality of the mutual fund family can be captured by return performance, then our results suggest that quality does not influence the advertising decision. On the other hand, if one assumes that quality of the fund family can be captured by the family's expenses as a proxy for service, then our results suggest that higher quality funds are the ones that advertise more heavily. Alternatively, one could interpret the coefficient on the expense ratio as suggesting that higher fee families can afford to expend more on advertising. it is also the case that larger families have a tendency to spend relatively more on advertising.

[40] The results do not change when current total net assets is used instead of lag assets.

Fund families with more fund objective classes do not advertise as much as do families with fewer objectives. This is consistent with the hypothesis that having more funds in different objective classes in and of itself provides more exposure for the fund family.

The relative amount of advertising is affected by the distribution channel decisions. Fund families with higher average load fees do not advertise as much as do fund families with lower average load fees. This result holds for both large and small families and is consistent with the hypothesis that advertising is directed more toward the retail investor than the financial advisers who would be receiving the commissions reflected in the load fees. Load funds rely more on the brokers and dealers, rather than advertising, to reach their investors.[41]

Relative advertising expenditures are increasing in families' 12b-1 fees, driven by the smaller families' 12b-1 fees. If we limit the analysis to large fund families, there is no relation between relative advertising and 12b-1 fees. One important implication of this result is that it points to a problem in studies that use 12b-1 fees to proxy for advertising expenditures. These proxies may be misleading, particularly for larger funds.

In sum, Table 8 shows that the fund family's other strategic decisions can affect the advertising decision.

7. Conclusions

Our interest in this paper is in the decisions a mutual fund family makes regarding the supply and demand of mutual fund products. We examine several strategic decisions on the family level, with a focus on the advertising decision. We examine the effects of these decisions on the family's net flows into its funds.

[41] This result is consistent with the results of Bergstresser, Chalmers and Tufano (2005) who find that…

We establish that the previously documented convexity in the relation between high past return performance and flows into the fund at the individual fund level (on an annual basis) also exists at the mutual fund family level (on a monthly basis) In contrast to evidence in Sirri and Tufano (1998) at the individual fund level, when using raw average family returns (without considering fund objective classes) our results also show a nonlinear relation for poor past return performance as well. Thus, past returns are a significant predictor of future family flows, but only for extreme relative returns. We also find, consistent with previous research on individual funds, that family level flows are related to load fees and 12b-1 fees.

The addition of advertising expenditures to this analysis does not significantly change the performance-flow relation, but advertising does affect fund flows. In fact, the form of the advertising-flow relation has a convexity at the upper end similar to that of the flow-performance relation, but is flat at the low end. That is, high relative levels of advertising are significantly related to high fund flows at the family level, while variations of relative levels of advertising within the low advertising group do not have a significant impact on flows to the family. Our results show that the relative amount of the expenditure has a nonlinear relation with fund flows. In fact, a simple linear specification of advertising expenditures does not identify the relationship between advertising and flows, coming in insignificant. Further, the increased flow from high relative levels of advertising is independent of the flow-performance relation.

We find that not advertising increases a family's average flow volatility (or average semivariance of flows), but the relative level of advertising does not affect this volatility. Whether the latter result is due to offsetting effects is a subject for future research with data sources that allow the differentiation between fund inflows and outflows. Average family flow volatility is related to the family's choice of distribution channels and its overall expenses.

Economic theory has suggested that high quality fund families should be the families that expend resources on advertising. If one assumes that the quality of fund families is reflected in their performance and in their services (proxied by expense ratios), then our results on this theory are mixed. No significant relation exists between prior year's returns and advertising expenditures, but we do find that the amount of advertising expenditures per dollar of assets is significantly related to a family's average expense ratio (excluding 12b-1 fees), which could be consistent with high quality funds advertising. These mixed results also require further study.

The results of our analyses indicate that previous proxies of marketing expenses do not reflect the entire picture as advertising expenditures have not been included. In particular, our findings that advertising increases in 12b-1 fees for small fund families but has no significant relation to 12b-1 fees for large fund families implies that studies that use 12b-1 fees to proxy for advertising expenditures are not capturing the true advertising expenditures and thus, the relation between fund flows and the advertising.

Overall, our results suggest that the fund's strategic decisions are important mechanisms through which mutual fund family management companies can affect their fund flows and consequent income. Specifically, we show that the relative level of advertising is an important strategic decision which has a nonlinear impact on the resulting flows to the fund family. Our work contributes to previous evidence on the other decisions by mutual fund family complexes, as well as to the literature that tries to understand the impact of increased visibility on investors' decisions

References

Bagwell, Kyle and Gary Ramey, 1994, Coordination economies, advertising and search behavior in retail markets, *American Economic Review* 84, 498-517.

Barber, Brad, and Terrence Odean, 2003, All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, *University of California-Davis and University of California-Berkeley working paper.*

Barber, Brad, Terry Odean, and Lu Zheng, 2005, Out of sight, out of mind – the effects of expenses on mutual fund flows, *Journal of Business,* forthcoming.

Baumol, William J., Steven M. Goldfeld, Lilli A. Gordon, and Michael F. Koehn, 1990, *The Economics of the Mutual Fund Markets: Competition versus Regulation*. Boston: Kluwer Academic Publisher.

Bergstresser, Daniel and James Poterba, 2002, Do aftertax returns affect mutual fund inflows? *Journal of Financial Economics*

Bergstresser, Daniel, John Chalmers and Peter Tufano, 2004, Assessing the costs and benefits of brokers in the mutual fund industry, working paper, Harvard University, University of Oregon.

Berk, Jonathan and Richard Green, 2004, Mutual fund flows and performance in rational markets. *Journal of Political Economy.*

Blume, Marshall, 1998, An anatomy of Morningstar ratings, *Financial Analysts Journal* 54, 1-9.

Capon, Noel, Gavan Fitzsimons, and Russ Prince, 1996, An individual level analysis of the mutual fund investment decision, *Journal of Financial Services Research* 10, 59-82.

Chauvin, Keith, and Mark Hirschey, 1993, Advertising, RD expenditures, and the market value of the firm, *Financial Management* 22, 128-?.

Chevalier, Judith and Glenn Ellison, 1997, Risk taking by mutual funds as a response to incentives. *Journal of Political Economy* 105, 1167-1200.

Christofferson, Susan, 2001, Why do money fund managers voluntarily waive their fees? *Journal of Finance* 56, 1117-1140.

Christofferson, Susan and David Musto, 2002, Demand curves and the pricing of money management, *Review of Financial Studies* 15, 1499-1524.

Chordia, Tarun, 1996, The structure of mutual fund charges. *Journal of Financial Economics* 41, 3-39.

Ciccotello, Conrad, Jason Greene, and Lori Walsh, 2003, No-transaction fee supermarket distribution of open-end mutual funds, Georgia State University working paper.

Collins, Sean and Phillip Mack, 1997, The optimal amount of assets under management in the mutual fund market, *Financial Analysts Journal* 67-73.

Cronqvist, Henrik, 2005, Advertising and portfolio choice, *Journal of Financial Economics,* forthcoming

Del Guercio, Diane and Tkac, Paula, 2002, The Determinants of the Flow of Funds of Managed Portfolios: Mutual Funds versus Pension Funds. *Journal of Financial and Quantitative Analysis* 37, 523-558.

Del Guercio, Diane and Tkac, Paula, 2002, Star power: the effect of Morningstar ratings on mutual fund flows, University of Oregon and Federal Reserve Bank of Atlanta working paper.

DeYoung, Robert and Evren Ors, 2005, Advertising and pricing at multiple output firms: Evidence from U.S. thrift institutions, Federal Reserve Bank of Chicago and HEC School of Management working paper.

Edelen, Roger, 1999. Investor flows and the assessed performance of open-end mutual funds. *Journal of Financial Economics* 53, 439-466.

Edelen, Roger, and Jerold Warner, 2001, Aggregate price effects of institutional trading: a study of mutual fund flow and market returns, *Journal of Financial Economics* 59,195 -220.

Elton, Edwin, Martin Gruber and Jeffrey Busse, 2004, Are investors rational? Choices among index funds, *Journal of Finance* 59, 261-288.

Elton, Edwin, Martin Gruber and Clifton Green, 2004, The impact of mutual fund family membership on investor risk, NYU and Emory working paper.

Fama, Eugene F., and James MacBeth, 1973, Risk, return, and equilibrium: Empirical tests, *Journal of Political Economy* 71, 607-636.

Fehle, Frank, Sergey Tsyplakov, and Vladimir Zdorovtsov, 2004, Can companies influence investor behavior through Advertising? Super Bowl Commercials and Stock Returns, University of South Carolina working paper.

Frieder, Laura and A. Subrahmanyam, 2005, Brand perceptions and the market for common stock, *Journal of Financial and Quantitative Analysis* 40.

Frye, Melissa, 2001, The performance of bank-managed mutual funds, *Journal of Financial Research* 24, 419-442.

Gaspar, J., M. Massa, and P. Matos, 2005, Favoritism in mutual fund families? Evidence of strategic cross-fund subsidization, *Journal of Finance*, forthcoming.

Goetzmann, William, and Nadav Peles, 1997, Cognitive dissonance and mutual fund investors, *Journal of Financial Research* 20, 145-158.

Golec, Joseph, 2003, Regulation and the rise in asset-based mutual fund management fees, *Journal of Financial Research* 26, 19-30.

Golec, Joseph, and Laura Starks, 2004, Performance fee contracts and mutual fund risk, *Journal of Financial Economics.*

Greene, Jason and Charles Hodges, 2002, The dilution impact of daily fund flows on open-end mutual funds, *Journal of Financial Economics* 65, 131-158.

Greer, D., 1971, Advertising and market competition, *Southern Economics Journal* 38, 19-32.

Gruber, Martin, 1996, Another puzzle: The growth in actively managed mutual funds, *Journal of Finance* 51, 783-810.

Grullon, Gustavo, George Kanatas, and James Weston, 2004, Advertising, breadth of ownership and liquidity, *Review of Financial Studies*.

Huang, Jennifer, Kelsey Wei, and Hong Yan, 2004, On the sensitivity of mutual fund flows to performance: theory and evidence, University of Texas and University of Binghamton working paper.

Ippolito, Richard A., 1992, Consumer reaction to measures of poor quality: Evidence from the mutual fund industry, *Journal of Law and Economics* 35, 45-70.

Jain, Prem C. and Joanna S. Wu, 2000, Truth in mutual fund advertising: Evidence on future performance and fund flows. *Journal of Finance* 55, 937-958.

Jayaraman, Narayanan, Ajay Khorana, and Edward Nelling, 2002, An analysis of the determinants and shareholder wealth effects of mutual fund mergers. *Journal of Finance* 57, 1521-1551.

Johnson, Woodrow T., 2004, Predictable investment horizons and wealth transfers among mutual fund shareholders, *Journal of Finance*.

Joshi, Amit and Dominique Hanssens, 2004, Advertsing spending and market capitalization, Marketing Science working paper.

Kempf, Alexander, and Stefan Ruenzi, 2005, Tournaments in mutual fund families. University of Cologne working paper.

Kempf, Alexander, and Stefan Ruenzi, 2004, Family matters: the performance flow relationship in the mutual fund industry. University of Cologne working paper.

Khorana, Ajay, and Henri Servaes,1999, The determinants of mutual fund starts, *Review of Financial Studies* 12,1043-1074.

Khorana, Ajay and Henri Servaes, 2003, An examination of competition in the mutual fund industry. Georgia Institute of Technology and London Business School working paper.

Khorana, Ajay and Henri Servaes, 1999, The determinants of mutual fund starts. *Review of Financial Studies* 12, 1043-1074.

Kihlstrom, Richard and Michael Riordan, 1984, Advertising as a signal, *Journal of Political Economy* 92, 427-450.

Latzko, David, 2001, Mutual fund expenses: an econometric investigation, Pennsylvania State University working paper.

Lee, Charles M. C., 1992, Earnings news and small traders, *Journal of Accounting and Economics*, 15, 265-302.

Leone, Robert P and Randall L Schultz, 1980, A study of marketing generalizations, *Journal of Marketing* 44, 10-18.

Lettau, Martin, 1997, Explaining the facts with adaptive agents: the case of mutual fund flows, *Journal of Economic Dynamics and Control* 21, 1117-1147.

Lynch, Anthony and David Musto, 2003, How investors interpret past fund returns, *Journal of Finance* 58, 2033-2058.

Mamaysky, Harry and Matthew Spiegel, 2002, A theory of mutual funds: Optimal fund objectives and industry organization, Yale University working paper.

Massa, Massimo, 2003, How do family strategies affect fund performance? When performance-maximization is not the only game in town, *Journal of Financial Economics* 67, 249-304.

Merton, Robert C., 1987, A simple model of capital market equilibrium with incomplete information, *Journal of Finance* 42, 483-510.

Milgrom, Paul, and Roberts, John, 1986, Price and advertising signals of product quality, *Journal of Political Economy* 94, 796-821.

Nanda, Vikram, M.P. Narayanan, and Vincent A. Warther, 2000, Liquidity, investment ability and mutual fund structure, *Journal of Financial Economics* 57, 417-443.

Nanda, Vikram, Z. Jay Wang, and Lu Zheng, 2004a, Family values and the star phenomenon, *Review of Financial Studies*, 667-698.

Nanda, Vikram, Z. Jay Wang, and Lu Zheng, 2004b, The ABCs of mutual funds: A natural experiment on fund flows and performance, University of Michigan working paper.

Nelson, Phillip, 1970, Information and consumer behavior, *Journal of Political Economy* 78, 311-329.

Nelson, Phillip, 1974, Advertising as information, *Journal of Political Economy* 81, 729-754.

Newey, W., and West, K. 1987. A simple, positive semidefinite heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica* 55: 703-708.

Ors, Evren, 2005, The role of advertising in commercial banks, HEC School of Management working paper.

Rakowski, David, 2003, Fund flow volatility and performance. Georgia State University working paper.

Reid, Brian and John Rea, 2003, Mutual funds distribution channels and distribution costs, *Investment Company Institute Perspective* 9, 1-19.

Reuter, Jonathan, and Eric Zitzewitz, 2004, The financial media, advertising, and mutual fund investor, University of Oregon and Stanford University working paper.

Sharpe, William, 1998, Morningstar's risk-adjusted ratings, *Financial Analysts Journal* 54, 21-33.

Siggelkow, Nicolaj, 2003, Why focus? A study of intraindustry focus effects, *Journal of Industrial Economics* 51, 121-150.

Sirri, Erik and Peter Tufano, 1998, Costly search and mutual fund flows, *Journal of Finance* 53, 1589-1622.

Tirole, J, 1995, *The Theory of Industrial Organization,* MIT Press, Cambridge Mass.

Tkac, Paula, 2004, Mutual funds: Temporary problem or permanent morass, Federal Reserve Bank of Atlanta working paper.

Wermers, Russ, 2003, Is money really smart? New evidence on the relation between mutual fund flows, manager behavior, and performance persistence, University of Maryland working paper.

Wilcox, Ronald, 2003, Bargain hunting or star gazing? Investors' preferences for stock mutual funds, *Journal of Business* 76, 645-663.

Table 1
Mutual Fund Family Characteristics

This table provides descriptive statistics on the sample and mutual fund family characteristics as of the first quarters of three years in the sample, 1992, 1996, and 2001. The table provides the number of mutual fund families in the sample along with their total assets. For the family characteristics, the table shows the total mutual fund assets under management, the aggregate monthly family flows as a percentage of assets, the percent of families with at least one 12b-1 fund share class, the percent of families with at least one fund share class with a front-end load fee, the average load fees across funds in the families, the average load fees across funds with loads in the family, the average expense ratio and the monthly advertising expenditures as a percentage of assets (in thousandths of a percent).

Date	Year		1992	1996	2001
Sample characteristics	Number of families		98	124	109
	Total assets ($billion)		935.37	1970.19	4224.97
Family characteristics	Total assets ($billion)	Mean	9.54	15.89	38.76
		S.D.	1.93	3.48	9.01
	Flows as a percentage of assets	Mean	4.35%	2.76%	1.79%
		S.D.	0.74%	0.44%	0.35%
	Percent with at least one 12b-1 fee fund share class		69.39%	77.42%	83.49%
	Percent with at least one front-end load fee fund share class		70.41%	76.61%	75.23%
	Average load fees	Mean	1.82%	1.53%	1.61%
		S.D.	0.22%	0.17%	0.18%
	Average load fees (load funds only)	Mean	4.46%	4.54%	4.98%
		S.D.	0.12%	0.08%	0.08%
	Average expense ratio	Mean	1.13%	1.21%	1.25%
		S.D.	0.06%	0.05%	0.05%
	Advertising monthly expenditures as a percentage of assets	Mean	3.35%	4.18%	2.26%
	(in thousandths of a percent)	S.D.	0.94%	0.79%	1.01%

Table 2

The Relation between Mutual Fund Family Flows, Previous Performance, and Strategic Decisions

This table provides the results of piecewise linear specifications of the mutual fund family flow with explanatory variables for that flow. Models 1 and 2 show the piecewise linear specifications with four kinks and two kinks, respectively. For the piecewise linear specifications, the family's value-weighted average return performance variable is broken into sub-variables that range from 0-.20 in the four-kink case (or 0-.33 in the two kink case) and the sum of which is equal to the original variable. The other variables are the lag flow from the previous month, the log of the total net assets (TNA), dummies for whether the following are in a family: star fund, front-end load fee, 12b-1 fee. Also included are load fees, 12b-1 fees and expense ratios (without 12b-1 fees) averaged across the funds in the family and the average is ranked against other families in the sample and average portfolio turnover. The models are run cross-sectionally each month from 1992-2001. The coefficients shown are the averages across the 114 months. The table also provides Newey-West (1987) t-statistics for the coefficients from the Fama-MacBeth (1973) aggregation technique and the average adjusted R-squareds from the regressions.

		Model 1		Model 2	
	Variable	**Mean coefficient**	**t-statistic**	**Mean coefficient**	**t-statistic**
	Intercept	0.004	1.00	0.006	1.23
Past returns	5th performance group	0.061	4.11 ***		
	4th performance group	-0.021	-1.56		
	3rd performance group	0.014	1.67 *	0.045	4.01 ***
	2nd performance group	0.005	0.73	0.002	0.66
	1st performance group	0.082	7.43 ***	0.089	8.02 ***
Strategic decisions	Log fund objectives offered	0.002	2.56 **	0.002	2.60 ***
	Dummy - front-end load fee	0.005	4.17 ***	0.006	4.30 ***
	Ranked average load fee	-0.008	-4.86 ***	-0.008	-4.91 ***
	Dummy - 12b-1 fees	-0.002	-1.01	-0.002	-0.99
	Ranked average 12b-1 fees	0.008	3.43 ***	0.008	3.79 ***
	Ranked average expense ratio	-0.007	-2.75 ***	-0.008	-2.95 ***
	Average turnover	-0.002	-2.07 **	-0.002	-2.21 **
Control variables	Lag Flow from previous month	0.067	2.26 **	0.067	2.28 **
	Log TNA	-0.001	-3.31 ***	-0.002	-3.19 ***
	Dummy - star fund in family	0.003	3.26 ***	0.003	3.26 ***
	Adj. R-squared	0.131		0.131	

Table 3

The Relation between Mutual Fund Family Flows, Adjusted Previous Performance, and Strategic Decisions

This table provides the results of piecewise linear specifications of the mutual fund family flow with explanatory variables for that flow. Models 1 and 2 show the piecewise linear specifications with four kinks and two kinks, respectively. In aggregating family return performance, each individual fund's performance is adjusted for the return performance in the fund's ICDI objective class. For the piecewise linear specifications, the family's value-weighted average excess return performance variable is broken into sub-variables that range from 0-.20 in the four-kink case (or 0-.33 in the two kink case) and the sum of which is equal to the original variable. The other variables are the lag flow from the previous month, the log of the total net assets (TNA), dummies for whether the following are in a family: star fund, front-end load fee, 12b-1 fee. Also included are load fees, 12b-1 fees and expense ratios (without 12b-1 fees) averaged across the funds in the family and the average is ranked against other families in the sample and average portfolio turnover. The models are run cross-sectionally each month from 1992-2001. The coefficients shown are the averages across the 114 months. The table also provides Newey-West (1987) t-statistics for the coefficients from the Fama-MacBeth (1973) aggregation technique and the average adjusted R-squareds from the regressions.

		Model 1		Model 2	
	Variable	**Mean coefficient**	**t-statistic**	**Mean coefficient**	**t-statistic**
	Intercept	0.004	0.83	0.006	1.12
Past returns	5th performance group	0.006	0.50		
	4th performance group	-0.018	-1.76 *		
	3rd performance group	0.012	1.37	-0.007	-0.62
	2nd performance group	0.007	0.89	0.002	0.99
	1st performance group	0.091	8.93 ***	0.096	10.50 ***
Strategic decisions	Log fund objectives offered	0.000	0.15	-0.001	-1.04
	Dummy - front-end load fee	0.004	2.56 **	0.004	2.73 **
	Ranked average load fee	-0.006	-3.16 ***	-0.005	-3.03 ***
	Dummy - 12b-1 fees	-0.002	-1.13	-0.002	-1.09
	Ranked average 12b-1 fees	0.007	3.35 ***	0.006	3.19 ***
	Ranked average expense ratio	-0.008	-3.04 ***	-0.008	-2.94 ***
	Average turnover	-0.002	-2.54 **	-0.002	-2.41 **
Control variables	Lag Flow from previous month	0.063	2.06 **	0.065	2.13 **
	Log TNA	0.000	0.73	0.000	0.36
	Dummy - star fund in family	0.001	1.53	0.002	1.76 *
	Adj. R-squared	0.127		0.126	

Table 4

The Relation between Aggregate Flows to Fund Families, Advertising Expenditures, and Aggregate Fund Performance

This table provides the results of times series regressions in which the dependent variable is the aggregate monthly flow to all fund families in our sample in Model 1 and aggregate monthly flows to the non-advertising fund families in Model 2. The independent variables are the lag flow to the family, the aggregate advertising expenditures across all funds, and the lag annual average performance across the fund families. The table also provides Newey-West (1987) t-statistics for the coefficients and the average adjusted R-squareds from the regressions.

	Model 1 All Families		Model 2 Non-advertising Families	
Variable	**Mean coefficient**	**t-statistic**	**Mean coefficient**	**t-statistic**
Intercept	0.002	1.06	0.000	0.15
Lag flow - previous month	-0.249	-2.88**	-0.231	-2.66**
Aggregate advertising expenditures	0.482	2.68**	0.177	1.93*
Lag performance – previous year	0.023	2.28**	0.042	2.62**
Adj. R-squared	0.093		0.081	

Table 5

The Relation between Mutual Fund Family Flows, Performance, and Strategic Decisions with Advertising Expenditures

This table provides the results of piecewise linear specifications of the family flow relation with explanatory variables including advertising expenditures. For comparison purposes, Model 1 shows the family flow relation without advertising from Table 2. Models 2 and 3 include the flow-performance advertising relation as well. For the piecewise linear specifications, the advertising variable is broken into sub-variables that range from 0-.33, and the sum of which is equal to the original variable. The other control variables are the lag flow from the previous month, the log of the total net assets (TNA), dummies for whether the following are in a family: star fund, front-end load fee, 12b-1 fee. Also included are load fees, 12b-1 fees and expense ratios (without 12b-1 fees) averaged across the funds in the family where the average is ranked against other families in the sample and average portfolio turnover. The models are run cross-sectionally each month from 1992-2001. The coefficients shown are the averages across the 114 month. The table also provides Newey-West (1987) t-statistics for the Fama-MacBeth (1973) coefficients and the average adjusted R-squareds from the regressions.

		Model 1		Model 2		Model 3	
	Variable	**Mean coefficient**	**t-statistic**	**Mean coefficient**	**t-statistic**	**Mean coefficient**	**t-statistic**
	Intercept	0.006	1.23	0.004	0.68	0.004	0.70
Ad variables	No advertising dummy			0.002	1.53	0.001	0.84
	Ranked advertising			0.002	1.08		
	Low advertising group					0.002	0.22
	Mid advertising group					-0.002	-0.71 **
	High advertising group					0.037	2.81 ***
Past returns	Low performance group	0.045	4.01 ***	0.046	4.04 ***	0.045	3.70 ***
	Mid performance group	0.002	0.66	0.001	0.41	0.001	0.36
	High performance group	0.089	8.02 ***	0.089	8.03 ***	0.090	8.02 ***
Strategic decisions	Log fund objectives offered	0.002	2.60 ***	0.002	2.56 **	0.002	2.82 ***
	Dummy - front-end load fee	0.006	4.30 ***	0.006	4.59 ***	0.006	4.56 ***
	Ranked average load fee	-0.008	-4.91 ***	-0.008	-4.41 ***	-0.008	-4.18 ***
	Dummy - 12b-1 fees	-0.002	-0.99	-0.002	-1.17	-0.002	-1.15
	Ranked average 12b-1 fees	0.008	3.79 ***	0.008	3.89 ***	0.008	3.45 ***
	Ranked average expense ratio	-0.008	-2.95 ***	-0.008	-3.10 ***	-0.008	-3.02 ***
	Average turnover	-0.002	-2.21 **	-0.002	-2.13 **	-0.002	-2.16 **
Control variables	Lag Flow from previous month	0.067	2.28 **	0.067	2.25 **	0.064	2.13 **
	Log TNA	-0.002	-3.19 ***	-0.001	-2.70 ***	-0.001	-2.61 ***
	Dummy - star fund in family	0.003	3.26 ***	0.003	3.10 ***	0.003	2.84 ***
	Adj. R-squared	0.131		0.128		0.123	

Table 6

The Effects of Advertising Expenditures on Mutual Fund Family Flow Volatility

This table provides the results of regressions of mutual fund family average flow volatility on advertising expenditures and control variables. Model 1 provides a linear specification for advertising in which advertising expenditures are ranked against other families in the sample. Model 2 provides a piecewise linear specification for advertising. The other strategic decision and control variables are the volatility from the previous year, the current flows into the family, the log of the total net assets (TNA), dummies for whether the following are in a family: star fund, front-end load fee, 12b-1 fee. Also included are load fees, 12b-1 fees and expense ratios (without 12b-1 fees) averaged across the funds in the family and the average is ranked against other families in the sample and average portfolio turnover. The models are run cross-sectionally each month from 1992-2001. The coefficients shown are the averages across the 114 months. The table also provides the Newey-West (1987) t-statistics for the Fama-MacBeth (1973) coefficients and the average adjusted R-squareds from the regressions.

		Model 1		Model 2	
	Variable	Mean coefficient	t-statistic	Mean coefficient	t-statistic
	Intercept	0.048	4.69 ***	0.049	4.54 ***
	Advertising	0.007	1.64		
	No advertising	0.009	3.98 ***	0.009	1.74 *
Ad variables	Low advertising group			0.009	0.28
	Mid advertising group			0.001	0.11
	High advertising group			0.051	1.57
	Low performance group	-0.074	-3.83 ***	-0.070	-3.57 ***
Past returns	Mid Performance group	-0.014	-3.46 ***	-0.014	-3.32 ***
	High performance group	0.035	1.62	0.039	1.82 *
	Log fund objectives offered	0.007	3.68 ***	0.014	11.03 ***
	Dummy - front-end load fee	0.013	10.48 ***	0.007	2.32 **
	Ranked average load fee	0.007	2.45 **	-0.015	-2.71 ***
Strategic decisions	Dummy - 12b-1 fees	-0.015	-2.94 ***	0.029	13.48 ***
	Ranked average 12b-1 fees	0.030	14.55 ***	0.004	0.83
	Ranked average expense ratio	0.003	0.82	-0.019	-3.60 ***
	Average turnover	-0.018	-3.60 ***	-0.002	-1.07
	Previous year flow volatility	-0.001	-0.74	0.466	17.70 ***
Control variables	Current family flows	0.465	17.60 ***	-0.034	-1.49
	Log lag TNA	-0.039	-1.75 *	-0.004	-3.22 ***
	Dummy - star fund in family	-0.003	-3.19 ***	0.008	3.97 ***
	Adj. R-squared	0.230		0.228	

Table 7
The Effects of Advertising Expenditures on the Semi-Variance of Mutual Fund Family Flows

This table provides the results of regressions of mutual fund family average flow semi-variance on advertising expenditures and control variables. Model 1 provides a linear specification for advertising in which advertising expenditures are ranked against other families in the sample. Model 2 provides a piecewise linear specification for advertising. The other strategic decision and control variables are the volatility from the previous year, the current flows into the family, the log of the total net assets (TNA), dummies for whether the following are in a family: star fund, front-end load fee, 12b-1 fee. Also included are load fees, 12b-1 fees and expense ratios (without 12b-1 fees) averaged across the funds in the family and the average is ranked against other families in the sample and average portfolio turnover. The models are run cross-sectionally each month from 1992-2001. The coefficients shown are the averages across the 114 months. The table also provides the Newey-West (1987) t-statistics for the Fama-MacBeth (1973) coefficients and the average adjusted R-squareds from the regressions.

	Variable	Model 1 Mean coefficient	Model 1 t-statistic	Model 2 Mean coefficient	Model 2 t-statistic
	Intercept	0.038	4.20 ***	0.034	3.57 ***
Ad variables	Advertising	0.004	1.20		
	No advertising	0.008	3.70 ***	0.012	2.96 ***
	Low advertising group			0.035	1.34
	Mid advertising group			-0.003	-0.53
	High advertising group			0.032	1.18
Past returns	Low performance group	-0.065	-4.66 ***	-0.061	-4.20 ***
	Mid Performance group	-0.016	-4.55 ***	-0.016	-4.42 ***
	High performance group	0.041	2.44 **	0.044	2.55 **
Strategic decisions	Log fund objectives offered	0.011	10.38 ***	0.012	11.19 ***
	Dummy - front-end load fee	0.008	3.24 ***	0.008	3.13 ***
	Ranked average load fee	-0.018	-3.43 ***	-0.017	-3.29 ***
	Dummy - 12b-1 fees	0.032	20.25 ***	0.031	18.72 ***
	Ranked average 12b-1 fees	0.002	0.45	0.002	0.57
	Ranked average expense ratio	-0.017	-3.99 ***	-0.017	-3.90 ***
	Average turnover	-0.001	-0.44	-0.001	-0.69
Control variables	Previous year flow semivariance	0.134	5.36 ***	0.133	5.30 ***
	Current family flows	-0.039	-1.93 *	-0.035	-1.69 *
	Log lag TNA	-0.002	-2.13 **	-0.002	-2.31 **
	Dummy - star fund in family	0.008	4.86 ***	0.008	5.14 ***
	Adj. R-squared	0.137		0.133	

Table 8
Determinants of Mutual Fund Family Annual Advertising Expenditures

This table provides the results of a regression of family advertising expenditures on a set of family characteristics. Model 1 presents the results for all families with a dummy variable if the family is a large family, defined as a family above the median in assets under management. Models 2 and 3 present the results when the regression is run separately for small and large families, respectively. The other control variables are the lag flow from the previous year, the lag volatility from the previous year, the log of the total net assets (TNA), dummies for whether the following are in a family: star fund, front-end load fee, 12b-1 fee. Also included are load fees, 12b-1 fees and expense ratios (without 12b-1 fees) averaged across the funds in the family and the average is ranked against other families in the sample and average portfolio turnover. The models are run cross-sectionally each year from 1992-2001. The coefficients shown are the averages across the 10 years. The table also provides the Newey-West (1987) t-statistics for the coefficients from the Fama-MacBeth (1973) aggregation technique and the average adjusted R-squareds from the regressions.

		All families Model 1		Small families Model 2		Large families Model 3	
	Variable	Mean coefficient	t-statistic	Mean coefficient	t-statistic	Mean coefficient	t-statistic
	Intercept	0.033	0.10	-0.286	-0.38	0.441	1.68 **
Past returns	Low performance	0.107	0.31	1.056	3.35 ***	-0.280	-1.01
	Mid performance	-0.065	-0.48	-0.059	-0.21	-0.144	-0.56
	High performance	-0.563	-0.76	-0.095	-0.10	0.396	0.36
Strategic decisions	Log fund objectives offered	-0.273	-3.61 ***	-0.185	-1.63	-0.004	-0.05
	Load dummy	0.135	1.56	0.069	0.48	0.074	0.61
	Ranked average load	-0.461	-3.30 ***	-0.436	-2.38 **	-0.183	-2.24
	12b-1 fees dummy	-0.033	-0.38	0.020	0.19	-0.091	-1.00
	ranked 12b-1	0.171	1.78 *	0.366	3.94 ***	-0.083	-1.10
	ranked expense ratio	0.361	2.80 ***	0.332	1.27	0.060	0.68
	Average turnover	0.090	1.96 *	0.064	1.02	0.087	1.62
Control variables	Previous year flow	0.201	1.53	0.031	0.21	0.296	1.16
	Previous year flow volatility	0.789	0.41	-0.353	-0.12	0.087	0.08
	Log lag TNA	0.065	1.98 **	0.051	0.58	-0.011	-0.34
	Dummy - star fund in family	0.042	1.09	0.093	1.70 *	-0.048	-0.51
	Size dummy	0.037	0.86				
	Adj. R-squared	0.286		0.134		0.197	

Figure 1
Total Net Assets

This figure shows how total assets under management for fund families in sample change through the sample period. Numbers are in billions of dollars.

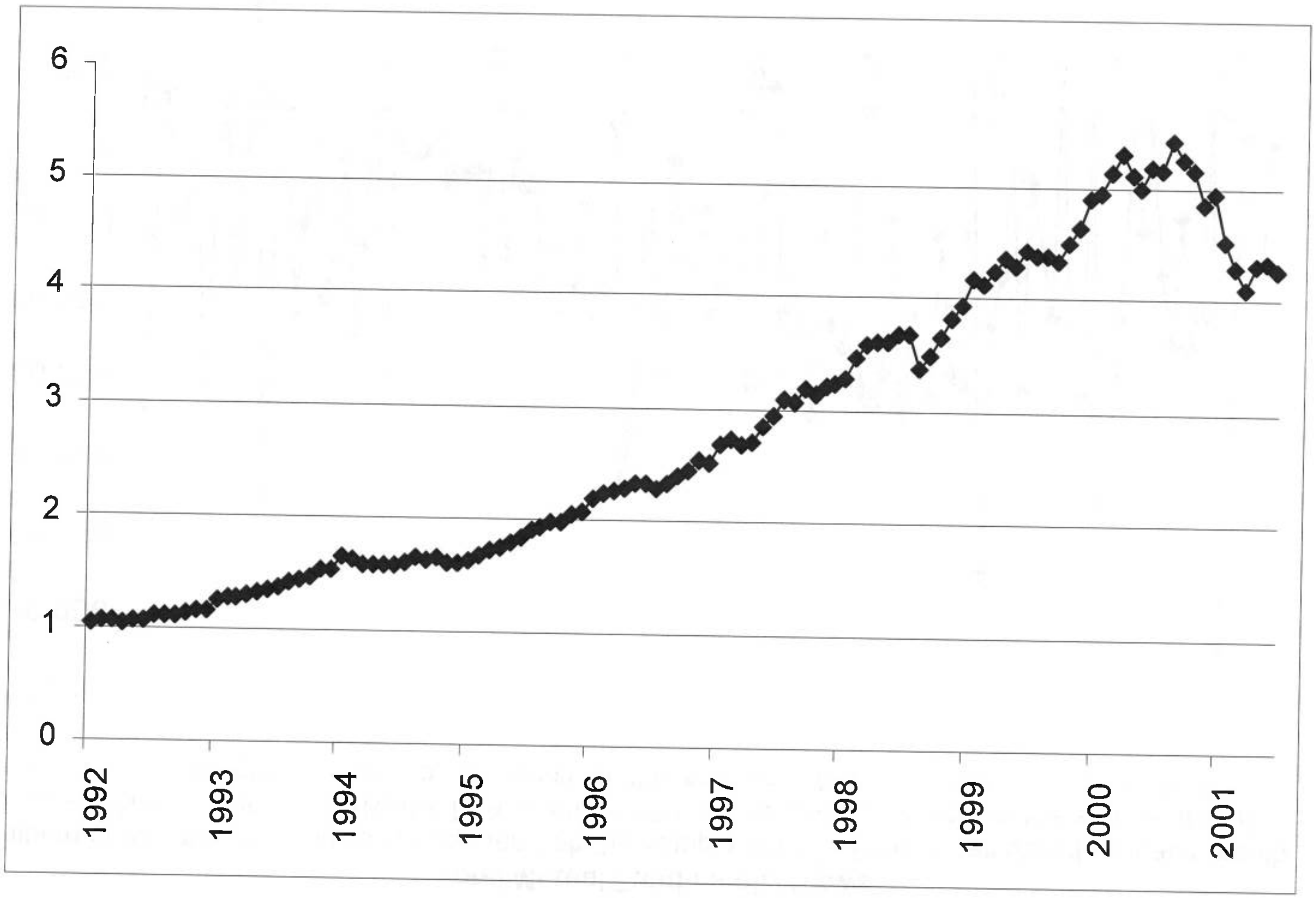


Figure 2
Total Mutual Fund Family Flows

This figure shows how total flows change through the sample period. Flow is calculated for each month as $Flow_t = \{TNA_t - (TNA_{t-1} * (1+R_t))\} / TNA_{t-1}$, where TNA is the total net assets in the sample and R is the market value weighted return for all funds in the sample. Numbers are in millions of dollars.

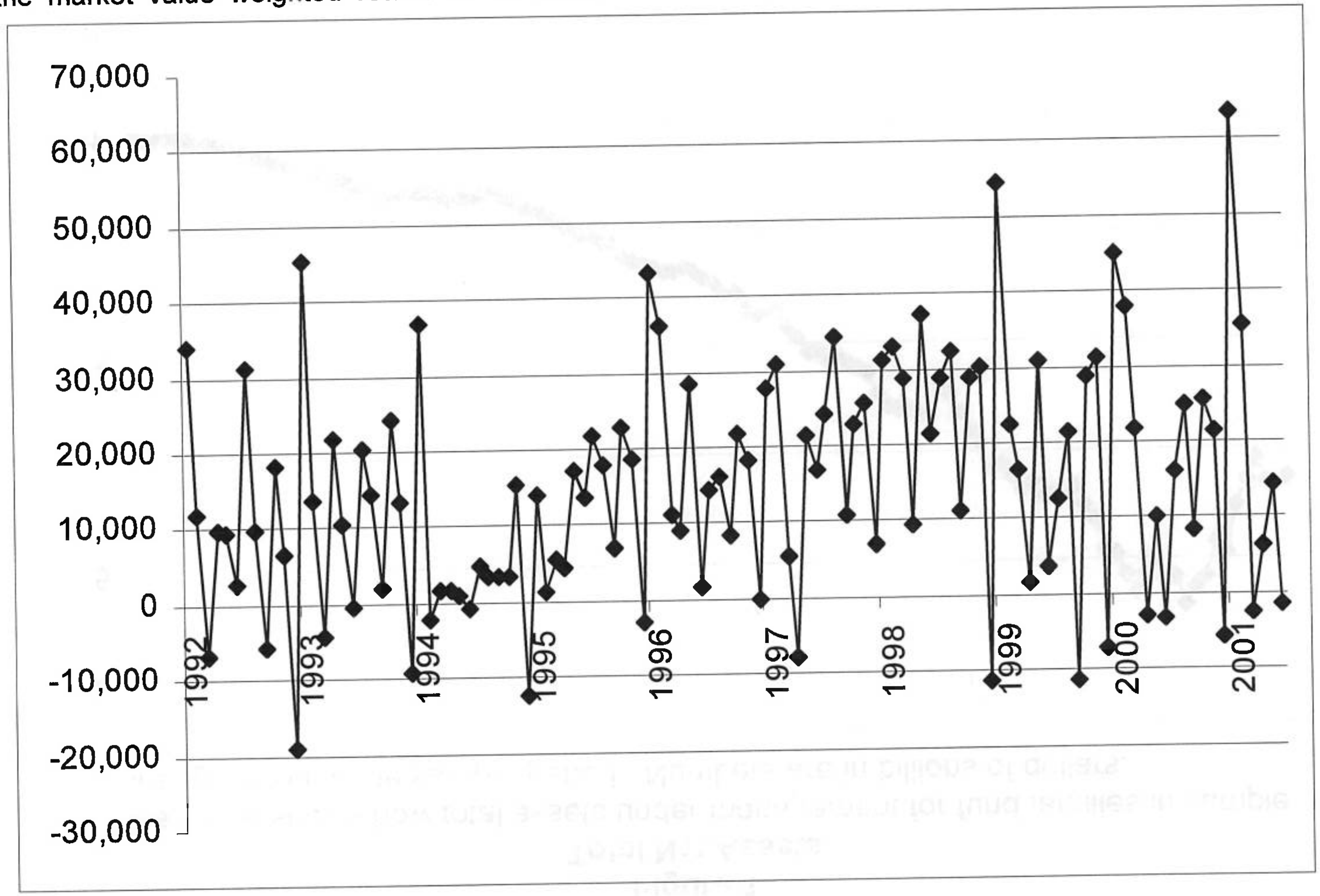


Figure 3
Dollar Advertising Expenditures

This figure shows how dollar advertising expenditures change through the sample period. Numbers are in millions of dollars.

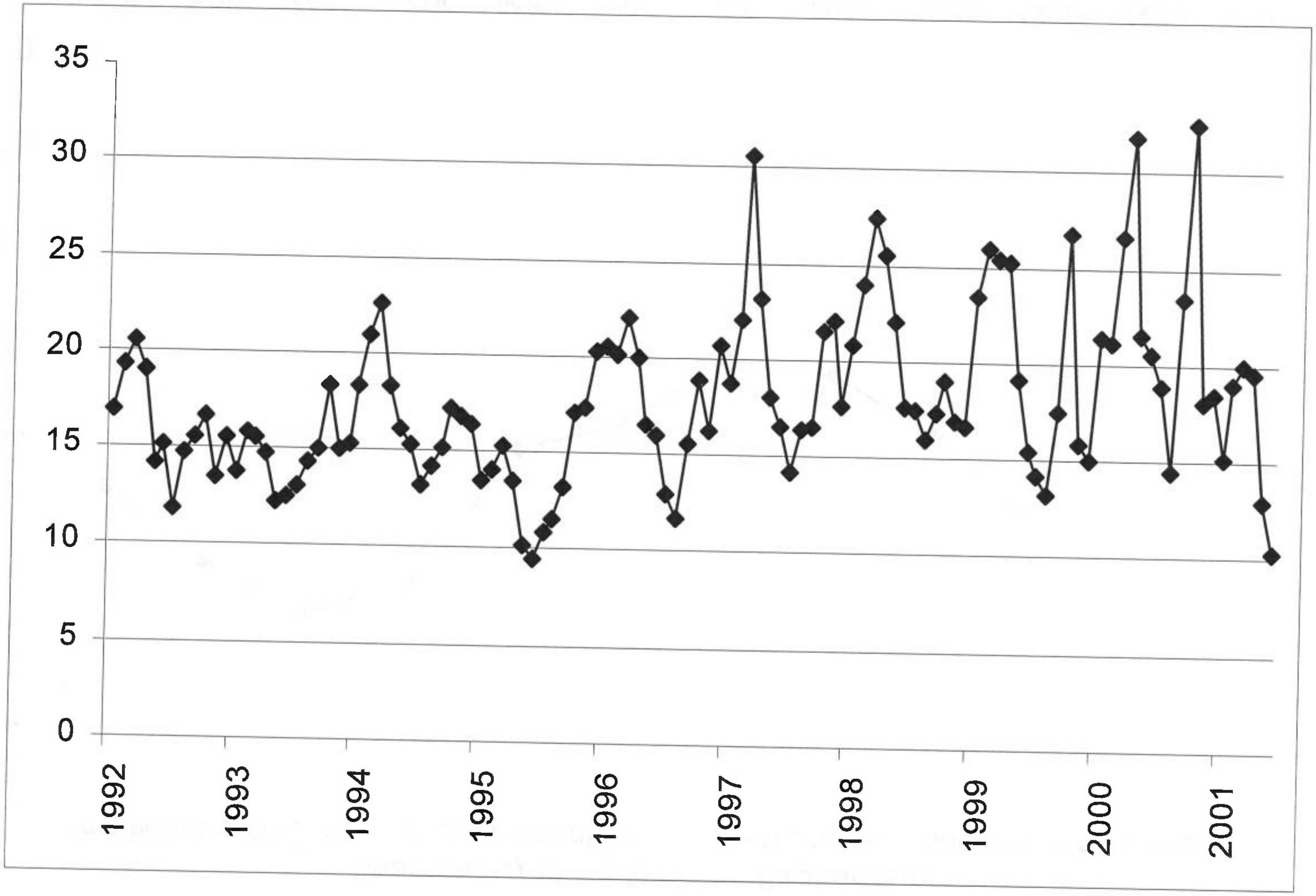


Figure 4
Seasonality in Advertising Expenditures
This figure shows the average percentage of advertising expenditures in each month.

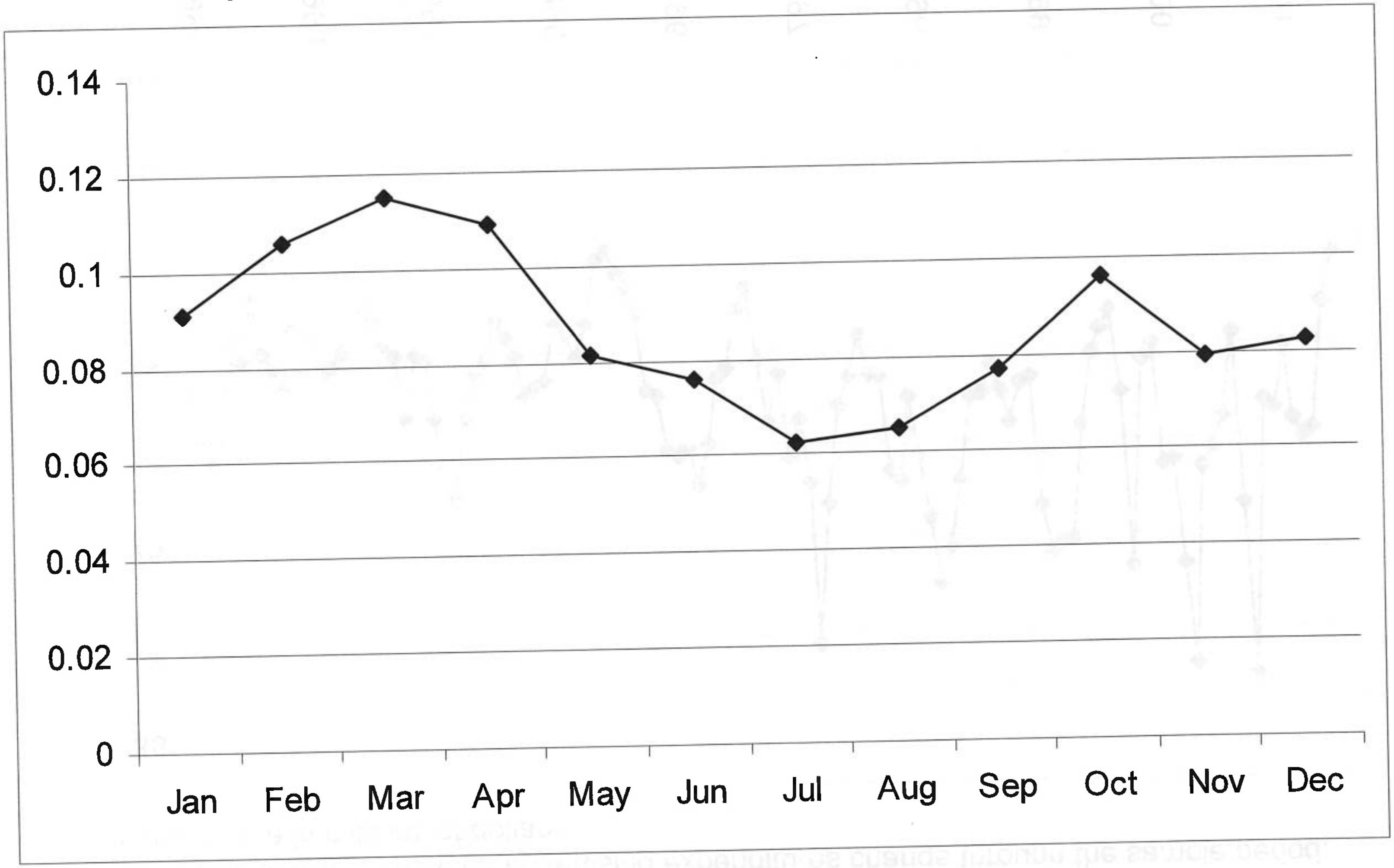
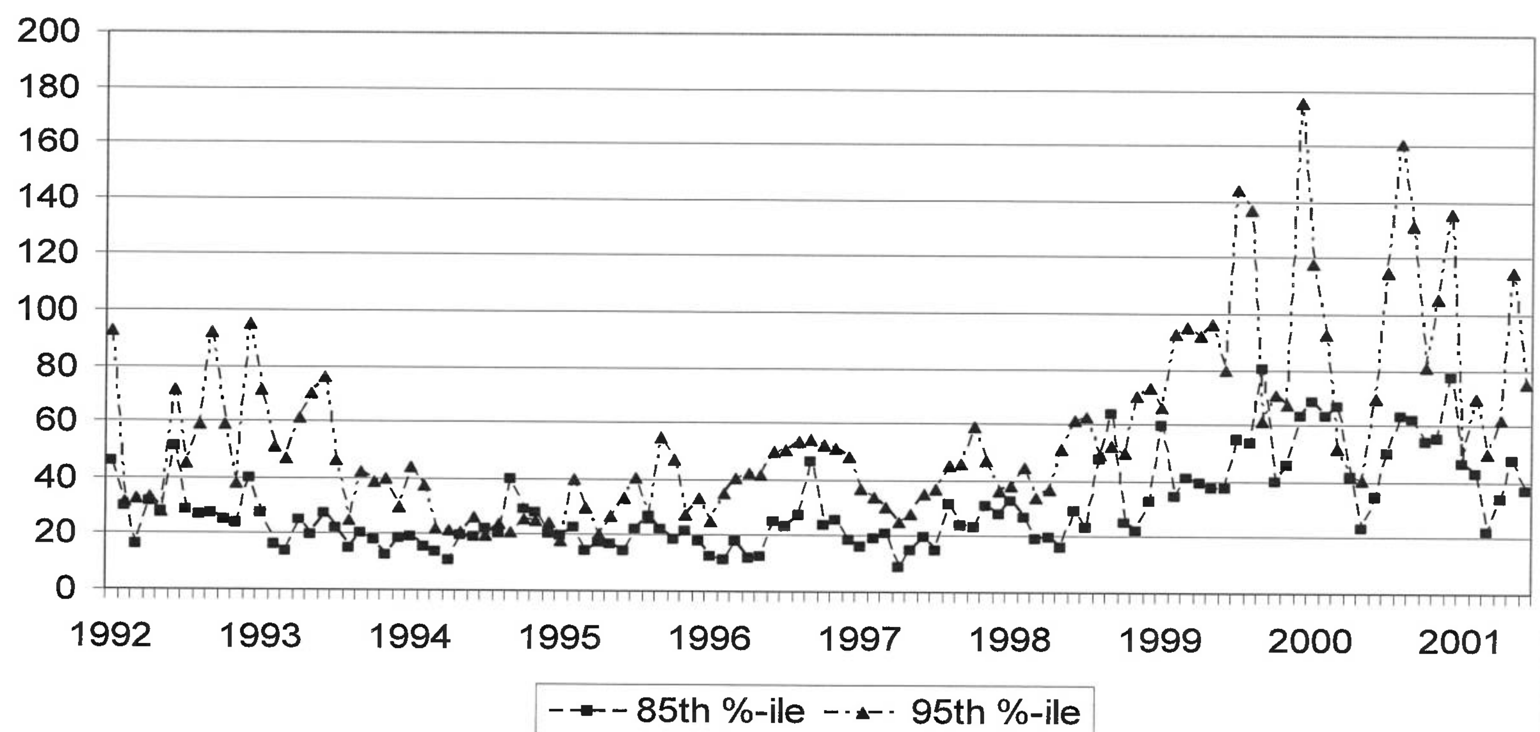


Figure 5
Estimated Total Return to Advertising

The figure shows the times series results for the 85th and 95th percentile advertisers at each point in time. At each point in time the figure shows the advertising expenditures of the fund family closest to the 85th (95th) percentile in advertising expenditures times the coefficient from that period's regression to generate the dollar flows to advertising. The dollar flows to advertising were then divided by the actual advertising cost to derive the returns to advertising.



47

HHS Public Access
Author manuscript
Manage Sci. Author manuscript; available in PMC 2016 November 01.

Published in final edited form as:
Manage Sci. 2015 November ; 61(11): 2603–2610. doi:10.1287/mnsc.2014.2010.

Aging and Financial Decision Making

Keith Gamble, PhD, **Patricia Boyle, PhD**, **Lei Yu, PhD**, and **David Bennett, MD**

Abstract

This study examines how cognitive changes associated with aging impact the financial decision making capability of older Americans. We find that a decrease in cognition is associated with a decrease in financial literacy. Decreases in episodic memory and visuospatial ability are associated with a decrease in numeracy, and a decrease in semantic memory is associated with a decrease in financial knowledge. A decrease in cognition also predicts a drop in self-confidence in general, but importantly, it is not associated with a drop in confidence in managing one's own finances. Participants experiencing decreases in cognition do show an increased likelihood of getting help with financial decisions; however, many participants experiencing significant drops in cognition still do not get help.

1. Introduction

There are concerns that financial decision making in the older population is compromised by the presence of cognitive decline in advanced age. These concerns combined with demographic changes resulting in an increasingly larger older population have sparked several recent studies on aging and financial decision making.[1] Recent studies find that older individuals are prone to worse financial decision making. For example, Korniotis and Kumar (2011) find that older investors exhibit less investment skill, and Agarwal et al. (2010) find that suboptimal credit behavior increases past age 53. Finke, Howe, and Huston (2011) suspect that financial decision making ability declines with age as financial literacy declines; indeed, they show that average financial literacy scores are lower by about 1% for each year after age 60. These existing studies indirectly examine the effects of cognitive aging on financial ability by comparing across individuals of different ages. Such comparisons confound the effect of cognitive decline with other differences, such as cohort effects. For example, Malmendier and Nagel (2011) find the cohort effect of early-life economic conditions on risk taking decades later. Direct measures of cognition collected repeatedly from the same individuals are needed in order to identify the effect of a decrease in cognition on financial ability.

We use longitudinal data from the Rush Memory and Aging Project, a large cohort study of aging, to identify, within individuals, the impact of decreases in cognition on financial literacy, financial confidence, and self-participation in financial decisions. In analyses restricted to persons without dementia based on a detailed clinical evaluation, we find that a decrease in cognition is a significant predictor of a decrease in financial literacy among

[1]For example, see the articles collected in Li, Ridderinkhof, and Samanez-Larkin (2011) and Samanez-Larkin (2011). Also, Samanez-Larkin and Knutson (2013) provide a recent summary of much of this work.

older Americans. Drops in cognition are associated with decreases in each of the components of financial literacy we measure, both numeracy and financial knowledge. We use our rich cognitive assessment including measures of five domains of cognition to identify which components of cognition are driving these changes in financial literacy. We find that decreases in episodic memory and visuospatial ability are associated with a decrease in numeracy while a decrease in semantic memory is associated with a decrease in financial knowledge.

Although participants experiencing decreased cognition also show declines in their financial literacy, these participants may not recognize or may be reluctant to admit to this decline in their financial capability. We find that a decrease in cognition predicts a drop in self-confidence in general, but importantly, it does not predict a decrease in confidence in managing one's own finances nor a decrease in confidence in one's financial knowledge. Similarly, Holland and Rabbitt (1992) find that individuals in their 70s do not rate their sensory abilities as poor any more so than individuals in their 50s despite significant declines in their measured ability. Importantly, they find that those older individuals who recognize their decline in sensory ability adjust their road-use behavior and have fewer accidents. Perhaps there is an analogy between driving and financial choices, and older Americans who have a drop in cognition would be more likely to take precautions in their financial decision making if made aware of the connection.

The detrimental effects of cognitive aging on the financial choices of older Americans can potentially be mitigated with help for financial decisions provided within or outside of the household. We find that individuals who experience a decrease in cognition are more likely to stop managing their own finances and pass on this responsibility to their spouse, and they are more likely to get financial help from outside their household. However, there are still many participants who are experiencing cognitive decreases who are not getting help with their financial decisions. Even among the participants experiencing statistically significant decreases in cognition, about half are not getting help with their financial decisions. While these participants are likely to benefit from trustworthy, knowledgeable advice, knowing who to trust in financial matters can be problematic.

2. Data Description and Construction of Measures

Our data come from the Rush Memory and Aging Project (MAP), an ongoing longitudinal study of aging (Bennett et al. (2012)). Since beginning in 1997, MAP has enrolled older participants from throughout the Chicago metropolitan area. Participants undergo yearly interviews and detailed clinical evaluations, including medical history, neurological, and neuropsychological examinations. The MAP data include demographic information for each participant, such as age, sex, and education. In 2010, a decision making assessment was added to MAP. The Institutional Review Board of Rush University Medical Center approved MAP and the decision making substudy.

We exclude data from the 71 participants who were diagnosed with dementia at the time of their first decision making assessment. For these participants even completing the decision making assessment is rare; only 19 of these participants provided answers to each of our outcomes of interest. Our conclusions are robust to including these participants, but we exclude them to avoid any selection bias due to the participants who could not complete the survey. Dementia is diagnosed in accordance with the standards set by the National Institute of Neurologic and Communicative Disorders and Stroke and the Alzheimer's Disease and Related Disorders Association (Bennett et al. (2005)). At the time of these analyses, 575 participants without dementia at the initial decision making assessment had completed at least two decision making assessments. Two decision making assessments are required to observe increases and decreases in decision making measures over time.

Since its beginning, MAP has collected yearly cognitive test scores for each participant. Cognition is assessed with 19 tests, which are listed in the appendix by the 5 cognitive domains assessed in the battery: episodic memory, perceptual speed, semantic memory, visuospatial ability, and working memory. Episodic memory captures the memory of specific events whereas semantic memory captures the knowledge of concepts. Working memory captures the ability to store and process transitory information. Perceptual speed involves the ability to process information quickly and make mental comparisons. Visuospatial ability involves understanding visual representations and the spatial relationships among objects. The raw scores of each of the 19 cognitive tests are converted to z-scores using the baseline mean and standard deviation of the entire MAP cohort on that test. These 19 z-scores are averaged to compute the global cognitive function score. The z-scores within each domain are averaged to compute each cognitive domain score. We calculate each participant's change in cognition from the first decision making assessment to the most recent decision making assessment.

We connect each participant's change in cognition to the concurrent change in a variety of measures related to financial decision making capability. The exact wording of each decision making question used in this study is provided in the appendix. The decision making questionnaire includes 16 standard financial literacy questions including 9 to test numeracy and 7 to test financial knowledge.[2] We measure financial literacy, numeracy, and knowledge by adding the number of correct answers in each category of questions. Participants may respond that they do not know the answer, and they can refuse to answer any question. Participants are aware of these possibilities: 26% of participants refused or said do not know to at least one financial literacy question during the decision making study. Typically, these participants refused or said do not know only rarely. Among the participants who used these options at least once, the average number is 1.65 refusals or do-not-knows per survey. These responses are treated the same as incorrect answers in this analysis.

Each financial knowledge question includes a follow up question immediately after to assess the participant's confidence in her answer to the preceding knowledge question using a four-point scale from extremely confident to not at all confident. We score an extremely confident answer as a 3, fairly confident as a 2, a little confident as a 1, and not at all

[2]The decision making questionnaire included two additional financial knowledge questions that were removed in this analysis because the wording of those questions varied from standard presentations. The results including the additional questions are consistent with the results presented.

confident as a 0. We measure each participant's confidence in her financial knowledge by summing the confidence scores to these 9 questions.

We also use two additional measures of confidence. We assess self-confidence using a single question that asks participants to report their general level of confidence on a ten-point scale with 1 indicating that they are not at all confident and 10 indicating that they are completely confident. We assess financial confidence with a single question that asks participants to report to what extent they agree with the statement: "I am good at managing day to day financial matters such as keeping up with checking accounts, credit cards, payments, and budgeting." Responses are reported on a seven-point scale from strongly agree indicating the highest level of financial confidence (6) to strongly disagree indicating the lowest level of financial confidence (0).

Participants are also asked who are primarily responsible for making their financial decisions. They are asked explicitly if they, their spouse, their child, or someone else is responsible, and they are asked to specify the relationship for a response that includes someone else. Thus, we can identify participants who make their own financial decisions, households who make their own financial decisions (participant or spouse), participants that get help with financial decisions (spouse or other person is specified, possibly in addition to self), and participants that get help from outside of the household (someone other than the participant or spouse is included as primarily responsible).

3. Cognition Change

3.1. Procedure for Cognition Change Sample

We use simple regressions to identify the effect of a change in cognition on these decision making variables. Each regression is of the following form:

$$\Delta y_i = a * \Delta\, Cognition_i + b + \varepsilon_i$$

In each regression the dependent variable is the change in the decision making variable (y) from the participant i's first decision making assessment to her most recent (Δy_i). When this dependent variable is binary, we use the logistic form for the regression.

The right-hand side includes participant i's change in cognitive function score ($\Delta Cognition_i$) and a constant. The coefficient of the first term (a) captures the effect of a one unit change in cognitive score. The error term is ε_i. The coefficient a captures the effect of both increases and decreases in cognition collectively. Since the focus of this study is on understanding the impact of decreases in cognition on financial decision making, we also run the previous regression using only the subset of participants who experience a decrease in cognition. For this subset the coefficient a only captures the association of decreases in cognition with the dependent variable.

3.2. Summary Statistics for Cognition Change Sample

Table 1 presents summary statistics for the 575 participants in the cognitive change sample. They are mostly female, well-educated, older Americans. The average age is 82.23 years, and only 23% are male. The participants average 15.11 years of education. About two-thirds (377 participants) of the sample experience a decrease in their global cognition z-score from their first decision making assessment to the most recent. The average decrease in measured cognition among this group is −0.29. Many participants (34%) increase their cognition score. The same questions are repeated each year, and participants benefit from the effect of practice. The average increase is smaller in size at 0.19.

Studies around the world find low levels of financial literacy (Lusardi and Mitchell (2011a)). Participants in this study perform similarly. Participants answer on average 11.20 of the 16 financial literacy questions correctly in their initial decision making assessment. They correctly answer the same percentage of numeracy questions and financial literacy questions on average (70%). This percentage does not change by much overall from first assessment to the most recent. Lusardi and Mitchell (2011b) analyze a three-question financial literacy module included in the 2004 Health and Retirement study. Two of their questions, one about inflation and one about compound interest, match questions used in our measure of literacy. They find that only 50% of respondents answered both the questions about inflation and compound interest correctly. In contrast, 65% of our respondents answered the same questions about inflation and compound interest both correctly.[3]

Participants display a high level of self-confidence: their self-confidence averages 7.17 on the 10 point scale with a 10 meaning completely confident. Confidence in managing finances is similarly high on average (4.98 out of 6), meaning that most participants agree with the statement that they are good at managing their day to day financial matters. Confidence in financial knowledge averages 14.77 out of a maximum of 21, which is a little higher than the score for a participant who indicates they are fairly confident for each question (14).

Consistent with their high confidence in their ability to manage finances and their high confidence in their financial knowledge, the vast majority of participants (88%) are primarily or jointly responsible for their financial decisions at the time of their first decision making assessment. About 41% get help with financial decisions, including from a spouse, child, or outside advisor. Just 25% get help with financial decisions from someone other than a spouse. Over time fewer participants make their own financial decisions and more get help. At time of the most recent decision making assessment, the percentage of participants making their own financial decisions dropped by 13%, and 11% more got help with financial decisions.

3.3. Cognition Changes and Literacy

In this subsection we examine the impact of decreases in cognition on financial literacy and its components (numeracy and financial knowledge). Table 2 presents results of six regressions following the form specified in the procedures section of this paper. Changes in cognition are associated with changes in financial literacy and its components. A one unit change in cognition is associated with a literacy change of 1.084, which comes from a 0.648

[3]The overlapping financial literacy questions are provided in the appendix as numeracy question 9 (inflation) and numeracy question 7 (compound interest).

change in numeracy and a 0.437 change in financial knowledge. Each association is statistically significant at the 1% level.

We rule out the possibility that the positive association between cognition changes and literacy changes in this regression could be driven by those participants with improvements in their cognition score improving their financial literacy scores as well and not from those with decreases in their cognition score getting worse on literacy. Since the impact of decreases in cognition is the focus of this study, we rule out the previous possibility by running the same regression only for the subset of participants whose global cognition score dropped. Decreases in cognition are associated with decreases in financial literacy and its components. A one unit decrease in cognition is associated with a financial literacy decrease of 1.237, which comes from a 0.765 decrease in numeracy and a 0.473 decrease in financial knowledge. The associations with literacy overall and numeracy are also statistically significant at the 1% level while the association with financial knowledge is statistically significant at the 5% level.

The size of these effects of cognitive changes on financial literacy are modest, but it is important to consider that the changes in cognition we are measuring during the decision making assessment period occurs over just two to three years. Individuals experiencing cognitive decreases are likely to experience further decreases over time. Thus, the impact of decreases in cognition on financial literacy is expected to accumulate over time.

3.4. Breakdown by Cognitive Domain

The association of drops in cognition with drops in financial literacy and its components can be separated into the five domains of cognition tested. Table 3 presents summary statistics for these five domain-specific cognitive measures. As with the global cognition score, participants' average score has dropped over time. The changes in these domain-specific cognitive measures are positively correlated; however, there is a lot of independent variation in these measures. The correlations range from a low of 6% between visuospatial ability and working memory to a high of 34% between episodic memory and semantic memory.

Table 3 also presents regression results for how decreases in these cognitive domain scores are associated with changes in the components of financial literacy. Numeracy changes are most strongly predicted by a drop in episodic memory; a one unit decrease in episodic memory is associated with a 0.725 decrease in numeracy, which is statistically significant at the 1% level. Numeracy changes are also associated with visuospatial ability (p=.03). Knowledge changes are most strongly predicted by a drop in semantic memory; a one unit decrease in semantic memory is associated with a 0.632 decrease in the participant's financial knowledge, which is statistically significant at the 5% level.

3.5. Cognition Changes and Confidence

We next examine the effect of changes in global cognition on a variety of confidence measures. First, we examine the effect of a decrease in cognition on general self-confidence. Table 4 shows that a one unit change in cognition is associated with a 0.416 change in self-confidence on a ten-point scale. This weak association in changes is driven by a strong association among the subset of those participants experiencing declining cognition. A one unit decrease in cognition is associated with a 0.968 decrease in self-confidence, which is statistically significant at the 5% level. However, we find a very different result for the effect of a decrease in cognition on one's confidence for managing financial matters. Neither changes in cognition or decreases in cognition are associated with changes in confidence in managing one's finances. Despite the drop in self-confidence associated with a decrease in cognition, participants who have a decrease in cognition do not reduce their confidence for managing their own finances.

Similarly, participants who experience a decrease in cognition do not significantly reduce their confidence in their financial knowledge. Although we find that a one unit change in cognition is associated with a 1.042 change in participants' confidence in their financial knowledge with statistical significance nearly at the 1% level, this result is not driven primarily by those with a decrease in cognition. In this case those increasing their cognition score are also increasing their confidence in their financial knowledge. Among those participants experiencing a decrease in cognition, there is only weak statistical significance in the association between decreases in cognition in the change in confidence in their financial knowledge with a p-value of 0.09.

In the previous subsection of this paper, we document a statistically significant finding that financial knowledge does drop with decreases in cognitive score; thus, these participants do not appear to recognize fully the detrimental effect of decreased cognition on their financial ability despite their decrease in self-confidence in general.

3.6. Cognitive Changes and Seeking Financial Help

Having shown that decreases in cognition are strongly associated with a decrease in financial literacy but not one's financial confidence, we now examine to what extent those participants who experience a decrease in their cognitive score get help with their financial decision making. Because the dependent variables in this subsection are binary, we alter our regression to the logistic form; otherwise, the explanatory variables are the same. Table 5 presents these logistic regression results. A one unit decrease in cognition results in an increase in the odds that a participant stops making her own financial decisions by $e^{1.098} - 1 = 203\%$. This relationship is statistically significant at the 1% level. Similarly, a one unit decrease in measured cognition results in an increase in the odds that both participant and spouse (a household) stop making their own financial decisions by $e^{1.290} - 1 = 263\%$ Again, this relationship is statistically significant at the 1% level.

Participants who experience a decrease in their cognition are more likely to obtain help with making financial decisions. A one unit decrease in measured cognition results in an increase in the odds that a participant obtained help for her financial decisions by $e^{0.864} - 1 = 137\%$. This result is statistically significant at the 5% level. It includes obtaining help from a spouse as well as anyone outside the household. Similarly, a one unit decrease in measured cognition results in an increase in the odds that a participant obtained help for her financial decisions from outside her household by $e^{0.878} - 1 = 141\%$, which is statistically significant at the 5% level. Typically, help from outside the household is provided by a son, a daughter, or a professional financial advisor.

Despite the strong association between decreases in cognition and seeking help with financial decisions, there are still many participants who experience significant declines in their cognition who are not getting help. We use each participant's complete history of cognitive scores, including those prior to the start of the Decision Making assessment, to determine the long-term cognitive trajectory of each individual. The number of annual cognition scores for participants in our sample ranges from 2 for the most recent enrollees to 15 for long-time participants. On average participants have 6.6 cognitive scores with median of 7. Thus, we have a long history of cognitive function scores to determine which participants are experiencing a decline in cognition during their time in MAP. For each participant we determine the slope of her cognitive ability by running a simple linear regression of cognition scores on age and a constant. There are 146 participants who have experienced both decreased cognition during the decision making assessment and a statistically significant cognitive decline during their participation in MAP. Of these 146 participants only about half (76) get help with their financial decision making.

4. Conclusion

We utilize the data from the Rush Memory and Aging Project and the Decision Making substudy to identify the detrimental impact of decreases in cognition associated with aging on the financial decision making ability of older Americans. We find that decreases in cognition are associated with decreases in financial literacy. We provide evidence that participants do not recognize this decrease. Despite showing significant drops in their self confidence in general, their confidence in their ability to manage their own finances and their confidence in their financial knowledge do not decrease with drops in measured cognition. Whether it is sought out or unsolicited, participants who experience a decrease in their cognitive score are more likely to obtain help with their financial decisions, though perhaps not as many get assistance as need it and bad advice may be a problem.

The importance of studying financial decision making in the older population has never been greater. Prior to 1980, retirees relied on a combination of employer-sponsored defined benefit pensions and Social Security for monthly income. For these retirees institutions shouldered the responsibility and the risk of investing contributions and managing payouts. Since 1980, many defined benefits plans have been replaced by defined contribution plans, which leave the responsibility of managing investments and withdrawals to the individual retiree. Poterba, Venti, and Wise (2008) document that in 2000, 87% of personal retirement contributions went to individual accounts with the largest proportion of these going to 401(k) accounts. The next generation of retirees will have the responsibility and risk of managing the money in these individual accounts sensibly. As the baby boom generation of Americans begins to retire, there will be an ever larger portion of the population shouldering this great financial responsibility of managing their own retirement wealth.

After the massive shift from defined benefit pensions to self-directed defined contribution retirement accounts, economists documented the many heuristics and biases of these new retirement savers (Benartzi and Thaler (2007)). Research also helped to reveal solutions such as automatic enrollment and default investment portfolios (Choi et al. 2004) that have greatly increased retirement savings. As this generation of workers begins to retire, we believe that research on the financial decision making of older Americans will be equally as important in revealing the heuristics, biases, and behaviors of this new generation of retirees. This information is essential to developing the innovations that will help them to maximize their well-being during this last period of their lives when many important and influential financial decisions are made.

Acknowledgments

This research was supported by the National Institute on Aging grants R01AG33678 and R01AG17917.

References

Agarwal, Sumit; Driscoll, John C.; Gabaix, Xavier; Laibson, David. The Age of Reason: Financial Decisions over the Life-Cycle and Implications for Regulation. Brookings Papers on Economic Activity. 2009; 2:51–117.

Benartzi, Shlomo; Thaler, Richard. Heuristics and Biases in Retirement Savings Behavior. Journal of Economic Perspectives. 2007; 21:81–104.

Bennett, David A.; Schneider, Julie A.; Buchman, Aron S.; Barnes, Lisa L.; Boyle, Patricia A.; Wilson, Robert S. Overview and Findings from the Rush Memory and Aging Project. Current Alzheimer Research. 2012; 9:646–663. [PubMed: 22471867]

Bennett, David A.; Schneider, Julie A.; Buchman, Aron S.; Mendes de Leon, Carlos; Bienais, Julia L.; Wilson, Robert S. The Rush Memory and Aging Project: Study Design and Baseline Characteristics of the Study Cohort. Neuroepidemiology. 2005; 25:163–175. [PubMed: 16103727]

Choi, James J.; Madrian, Brigitte C.; Laibson, David. Plan Design and 401(k) Savings Outcomes. National Tax Journal. 2004; 62:275–298.

Finke, Michael S.; Howe, John; Huston, Sandra J. Old Age and the Decline in Financial Literacy. 2011 working paper, Available at SSRN: http://ssrn.com/abstract=1948627.

Holland, Carol A.; Patrick, M.; Rabbitt, A. People's Awareness of Their Age-Related Sensory and Cognitive Deficits and the Implications for Road Safety. Applied Cognitive Psychology. 1992; 6:217–231.

Korniotis, George M.; Kumar, Alok. Do Older Investors Make Better Investment Decisions? Review of Economics and Statistics. 2011; 93:244–265.

Li, Shu-Chen; Richard Ridderinkhof, K.; Samanez-Larkin, Gregory R., editors. Frontiers of Neuroscience. Volume 5. Decision Making Across the Life Span. 2011.

Lusardi, Annamaria; Mitchell, Olivia S. Financial Literacy Around the World: An Overview. Journal of Pension Economics and Finance. 2011; 10:497–508.

Lusardi, Annamaria; Mitchell, Olivia S. Financial Literacy and Planning: Implications for Retirement Wellbeing. 2011 NBER working paper 17078.

Malmendier, Ulrike; Nagel, Stefan. Depression Babies: Do Macroeconomic Experiences Affect Risk Taking? Quarterly Journal of Economics. 2010; 126:373–416.

Poterba, James; Venti, Steven; Wise, David. The Changing Landscape of Pensions in the United States. In: Lusardi, Annamaria, editor. Overcoming the Saving Slump: How to Increase the Effectiveness of Financial Education and Savings Programs. Chicago, IL: University of Chicago Press; 2008.

Samanez-Larkin, Gregory R., editor. Decision Making Over the Life Span. Vol. 1235. New York: Annals of the New York Academy of Sciences; 2011.

Samanez-Larkin, Gregory R.; Knutson, Brian. Reward Processing and Risky Decision Making in the Aging Brain. In: Reyna, Valerie F.; Zayas, Vivian, editors. The Neuroscience of Risky Decision Making. Washington, DC: American Psychological Association; 2014. forthcoming

Appendix

Cognition Tests

The global cognition score is calculated by converting raw scores on each of the 19 tests listed below to z scores, using the mean and standard deviation from the full cohort at baseline, and then averaging the z scores to produce the composite measure. The composite measure of each cognitive domain is calculated similarly using only the tests in that domain.

Episodic Memory

1. Logical memory (immediate) Story A from the Logical Memory subset of the Wechsler Memory Scale-Revised;
2. Logical memory (delayed) Story A from the Logical Memory subset of the Wechsler Memory Scale-Revised;
3. CERAD Word list recall (immediate)
4. CERAD Word list recall (delayed)
5. CERAD Word list recognition
6. East Boston Story (immediate)
7. East Boston Story (delayed)

Perceptual Speed

1. Oral version of the Symbol Digit Modalities Test
2. Number Comparison
3. 2 indices from a modified version of the Stroop Neuropsychological Screening test

Semantic Memory

1. Verbal fluency from CERAD;
2. 15 item version of the Boston Naming Test
3. 15-item reading test

Visuospatial Ability

1. 15-item version of Judgment of Line Orientation
2. 16-item version of Standard Progressive Matrices

Working Memory

1. Digit Span subtests-forward of the Wechsler Memory Scale-Revised
2. Digit Span subtests-backward of the Wechsler Memory Scale-Revised
3. Digit Ordering

Survey Questions

Numeracy Questions

1. Which of these percentages represents the biggest risk of getting a disease? 1%, 10%, 5%
2. A store is offering 15% off a television that is normally priced at $1000. How much money would you save on the TV during this sale? $15, $150, $1500
3. If a television set is on sale for $899, which is $200 off its normal price, what is the normal price? $699, $1099, $1299
4. If 5 people all have the winning numbers in the lottery and the prize is $2 million, how much will each of them receive? $200,000; $400,000; $600,000
5. If the chance of getting a disease is 10 percent, how many people out of 1,000 would be expected to get the disease? 100, 10, 90, 900
6. In a sale, a shop is selling all items at half price. Before the sale, a sofa costs $300. How much will it cost in the sale? $150, $600, $900
7. Suppose you had $100 in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow: more than $102, exactly $102, or less than $102?
8. Again, suppose you had $100 in a savings account and the interest rate was 2% per year. After 5 years, how much do you think you would have in the account if you left the money to grow: more than $110, exactly $110, or less than $110?
9. Imagine that the interest rate on your savings account is 1% per year and inflation is 2% per year. After 1 year, will you be able to buy more than, exactly the same as, or less than today with the money in your account?

Financial Knowledge and Confidence Questions for Overconfidence Measure

Note: Each financial knowledge question is followed by the same confidence question below.

How confident are you that you answered that question correctly?

extremely confident, fairly confident, a little confident, not at all confident

1. What do the initials FDIC stand for?
2. What does the FDIC do?

   approves new drugs for clinical use, protects the funds people or depositors place in banks and savings institutions, underwrites mortgages and other loans

3. When interest rates go up, what do bond prices do: go down, go up, or stay the same?

4. True or False. An older person with $100,000 to invest should hold riskier financial investments than a younger person with $100,000 to invest.

5. True or False. Using money in a bank account to pay off credit card debt is usually wise.

6. True or False. To make money in the stock market, you have to buy and sell stocks often.

7. True or False. Stocks and mutual funds generally produce higher average returns above inflation compared to fixed-income investments such as bonds.

Self-confidence Question

Using a 1–10 point rating scale, where 1 indicates that you are not at all confident and 10 indicates that you are completely confident, how would you rate your general level of confidence?

Financial Confidence Question

Please give a number between 1 and 7 indicating the degree to which you agree or disagree with this statement, with 1 being strongly agree and 7 strongly disagree. I am good at managing day to day financial matters such as keeping up with checking accounts, credit cards, payments, and budgeting.

Question about Who Makes Financial Decisions

Currently, who is (are) primarily responsible for making your financial decisions: you, your spouse/partner, your child, or someone else? If someone else, please specify the other person.

Table 1

Summary Statistics for Cognition Change Sample

	All Participants		Decrease in Cognition	
Participants	575		377	
Male	23%		22%	
Age	82.23 (7.36)		83.17 (7.19)	
Education	15.11 (2.86)		15.21 (2.93)	
	Initial Level	**Change**	**Initial Level**	**Change**
Cognition	0.22 (0.54)	−0.13 (0.35)	0.22 (0.55)	−0.29 (0.30)
Financial Literacy	11.20 (2.30)	−0.15 (2.10)	11.08 (2.36)	−0.34 (2.21)
Numeracy	6.32 (1.35)	−0.11 (1.54)	6.26 (1.37)	−0.24 (1.58)
Financial Knowledge	4.88 (1.47)	−0.04 (1.32)	4.82 (1.49)	−0.10 (1.40)
Self-Confidence	7.17 (1.83)	0.08 (1.94)	7.17 (1.88)	0.06 (2.04)
Confidence in Managing Finances	4.98 (1.38)	−0.03 (1.33)	4.88 (1.45)	−0.12 (1.52)
Confidence in Financial Knowledge	14.77 (4.33)	0.02 (3.56)	14.52 (4.41)	−0.19 (3.60)
Participant Makes Financial Decisions	88%	−13%	87%	−16%
Household Makes Financial Decisions	91%	−10%	91%	−12%
Gets Help with Financial Decisions	41%	12%	45%	12%
Gets Help Outside of Household	25%	11%	29%	12%

This table presents summary statistics for the whole sample and the subsample of participants who experienced a decrease in cognition score during the Decision Making assessment. Age and Education are stated in years. Cognition is a z-score scaled to all participants in the Memory and Aging Project at baseline. Values are reported as means (standard deviation) or percentages.

Table 2

Decrease in Cognition and Literacy

	Literacy Change			Numeracy Change			Knowledge Change		
	Estimate	SE	p-value	Estimate	SE	p-value	Estimate	SE	p-value
Cognition Change	1.084***	0.248	0.00	0.648***	0.183	0.00	0.437***	0.158	0.01
Intercept	−0.012	0.092	0.90	−0.031	0.068	0.64	0.020	0.058	0.74
Cognition Decrease	1.237***	0.375	0.00	0.765***	0.269	0.00	0.473**	0.240	0.05
Intercept	0.021	0.157	0.89	−0.017	0.113	0.88	0.038	0.101	0.71

This table presents regression results of the effect of cognitive changes on financial literacy, and separately on its components, numeracy and knowledge. Cognition Change provides the association between participants' cognition changes and their changes in the specified outcome variable. Cognition Decrease provides the previous association only among the subset of participants who experience a decrease in cognition.

*, **, and *** indicate statistically significant at the 10%, 5%, and 1% level respectively

Table 3

The Components of Global Cognition

	Decrease in Cognition	
	Initial Level	Change
Episodic Memory	0.32 (0.74)	−0.30 (0.42)
Perceptual Speed	0.09 (0.78)	−0.31 (0.46)
Semantic Memory	0.22 (0.62)	−0.21 (0.39)
Visuospatial Ability	0.24 (0.74)	−0.26 (0.61)
Working Memory	0.15 (0.74)	−0.22 (0.49)

	Numeracy Change			Knowledge Change		
	Estimate	SE	p-value	Estimate	SE	p-value
Episodic Memory Decrease	0.725***	0.242	0.00	0.294	0.213	0.17
Intercept	0.173	0.130	0.18	0.024	0.114	0.83
Perceptual Speed Decrease	0.369*	0.219	0.09	0.212	0.191	0.27
Intercept	−0.022	0.122	0.86	0.013	0.107	0.90
Semantic Memory Decrease	0.404	0.300	0.17	0.632**	0.246	0.01
Intercept	0.065	0.134	0.63	0.203*	0.112	0.07
Visuospatial Ability Decrease	0.472**	0.212	0.03	0.024	0.194	0.90
Intercept	0.136	0.143	0.34	−0.060	0.131	0.65
Working Memory Decrease	0.333	0.254	0.19	0.153	0.211	0.47
Intercept	−0.017	0.145	0.68	−0.010	0.120	0.42

This table presents summary statistics and regression results for the effect of cognitive changes within each of the five domains on the components of financial literacy, numeracy and knowledge. Cognition Decrease provides the association of the change in cognitive domain score with the change in the specified component of financial literacy only among the subset of participants who experience a decrease in cognition.

*, **, and *** indicate statistically significant at the 10%, 5%, and 1% level respectively

Table 4

Decrease in Cognition and Confidence

	Change in Self-Confidence			Change in Confidence in Managing Finances		
		SE	p-value	Estimate	SE	p-value
Cognition Change	0.416*	0.234	0.08	0.163	0.168	0.33
Intercept	0.131	0.086	0.13	−0.094	0.061	0.13
Cognition Decrease	0.968***	0.350	0.01	0.098	0.267	0.72
Intercept	0.342**	0.147	0.02	−0.092	0.110	0.40

	Change in Confidence in Financial Knowledge		
	Estimate	SE	p-value
Cognition Change	1.042**	0.426	0.01
Intercept	0.153	0.157	0.33
Cognition Decrease	1.062*	0.616	0.09
Intercept	0.126	0.259	0.63

This table presents regression results of the effect of cognition changes on three different measures of confidence. Cognition Change and Cognition Decrease are defined as in Table 2.

*, **, and *** indicate statistically significant at the 10%, 5%, and 1% level respectively

Table 5

Decrease in Cognition and Seeking Financial Help

	Participant Stopped			Household Stopped		
	Estimate	SE	p-value	Estimate	SE	p-value
Cognition Change	-1.123^{***}	0.323	0.00	-1.119^{***}	0.347	0.00
Intercept	-2.113^{***}	0.145	0.00	-2.414^{***}	0.164	0.00
Cognition Decrease	-1.098^{***}	0.406	0.01	-1.290^{***}	0.432	0.00
Intercept	-2.043^{***}	0.203	0.00	-2.462^{***}	0.232	0.00
	Obtained Help			**Obtained Help Outside Household**		
	Estimate	**SE**	**p-value**	**Estimate**	**SE**	**p-value**
Cognition Change	-0.658^{**}	0.331	0.05	-0.787^{**}	0.341	0.02
Intercept	-2.065^{***}	0.142	0.00	-2.227^{***}	0.151	0.00
Cognition Decrease	-0.864^{**}	0.427	0.04	-0.878^{**}	0.434	0.04
Intercept	-2.205^{***}	0.217	0.00	-2.284^{***}	0.223	0.00

This table presents logistic regression results of the effect of cognitive changes on participants' participation in their own financial decisions. Cognition Change and Cognition Decrease are defined as in Table 2.

*, **, and *** indicate statistically significant at the 10%, 5%, and 1% level respectively

48

RAND Report

Potential Economic Effects on Individual Retirement Account Markets and Investors of DOL's Proposed Rule Concerning the Definition of a 'Fiduciary'

STEVEN GARBER, JEREMY BURKE, ANGELA A. HUNG, AND ERIC TALLEY

RAND LABOR AND POPULATION

RR-1009-DOL
February 2015
Prepared for the Department of Labor

For more information on this publication, visit www.rand.org/t/RR1009

Published by the RAND Corporation, Santa Monica, Calif.

Preface

In October 2010 the U.S. Department of Labor (DOL), Employee Benefits Security Administration (EBSA) proposed a new rule, 29 *Code of Federal Regulation* (CFR) Part 2510, "Definition of the Term 'Fiduciary'—Proposed Rule" (the Proposed Rule; U.S. DOL, EBSA, 2010) that would redefine the circumstances under which organizations and individuals are considered to be "fiduciaries" by reason of giving investment advice to an employee benefit plan, a plan participant, or an individual retirement account (IRA) holder. A key objective of the Proposed Rule is to reduce the incidence of self-dealing by broker-dealers and their associated representatives in the course of providing investment advice.

Among the comments received in response to the Proposed Rule were two reports from consulting firms that focused on potential unintended negative consequences for individuals who save for retirement through traditional and Roth IRAs. EBSA's Office of Policy and Research asked RAND to review and analyze economic issues raised by those two comments and provide its thoughts on the validity and policy implications of predictions made therein. This paper is RAND's response to that request.

The work reported here was undertaken in 2012 and was sponsored by the DOL. The report should be of interest to DOL staff; staff of other federal agencies that have regulatory responsibilities related to financial markets, such as the U.S. Securities and Exchange Commission; broker-dealers who provide advisory services related to IRAs; and economists and policy analysts with interests related to potential effects of conflicts of interest for financial advisors and potential policy responses.

This research was undertaken within the Center for Financial and Economic Decision Making (CFED). The mission of CFED, a part of RAND's Labor and Population research division, is to understand how people in the United States and around the world collect and think about financial information and how successfully they match their financial decisions to their interests and goals. CFED's researchers are dedicated to finding solutions that can improve the decisionmaking that affects the financial well-being of individuals, families, and nations. RAND Labor and Population has built an international reputation for conducting objective, high-quality, empirical research to support and improve policies and organizations around the world.

For more information about CFED, contact the Director, Angela Hung, by email at Angela_Hung@rand.org; by phone at 310-393-0411; or by mail at the RAND Corporation, 1776 Main Street, P.O. Box 2138, Santa Monica, CA, 90407-2138. More information about RAND is available at www.rand.org.

Table of Contents

Tables

Summary

In October 2010, the Employee Benefits Security Administration (EBSA) of the U.S. Department of Labor (DOL) published in the *Federal Register* a Proposed Rule that would broaden the circumstances under which organizations and individuals are considered to be "fiduciaries" by reason of giving investment advice to an employee benefit plan, plan participant, or individual retirement account (IRA) holder. The primary objective of the rule is to protect retirement investors from "self-dealing"—namely investment advisors acting in their own interests to the detriment of their clients' interests—by financial services firms such as broker-dealers and individuals who offer advice to retirement investors.

Two of the comments on the DOL proposal focused on potential undesirable effects of adopting the rule on investors who save for retirement through traditional or Roth IRAs. We analyze predictions made in these two comments, relying on standard principles of microeconomics and available empirical literature. We also present and illustrate a framework for predicting the effects of adoption of the Proposed Rule on the well-being of IRA investors.

One of the two comments argues that—despite the existence of financial incentives for advisors to self-deal—advisors rarely, if ever, act on these conflicts of interest. More specifically, that comment argues that incentives stemming from competition among broker-dealers for clients and the desire of these firms to have good reputations prevent self-dealing. In turn, they argue that—because self-dealing is currently rare or nonexistent—the adoption of the Proposed Rule would offer essentially no social benefit. We review these arguments and find them unconvincing. Moreover, we discuss empirical literature, which the commenters do not discuss, that does support the conclusion that self-dealing is, in fact, sufficiently widespread that substantial social benefits could result from reducing the prevalence of self-dealing. We then consider the potential effects of adopting the rule, including effects on the behavior of broker-dealers and effects on the well-being of IRA investors.

If the Proposed Rule were adopted, some forms of advisor compensation linked to particular investment transactions on behalf of IRA investors—a full list of which has yet to be determined—would be newly prohibited. For example, adoption of the rule would prohibit some forms of payments from third parties (such as sellers of financial products to broker-dealers) that are conditional on their clients investing in ways that benefit these third parties. Moreover, adoption of the Proposed Rule could increase some costs to broker-dealers, such as costs of monitoring advisor behavior for compliance with the DOL rule and costs of additional lawsuits and liability insurance. Eliminating some current sources of revenues and increasing their costs may cause broker-dealers to respond with efforts to increase revenues from allowed sources, reduce costs to serving the IRA market, or both. Costs might be reduced, for example, by reducing the

"intensity" of advice provided to IRA clients, which we define as the average time and effort advisors devote to serving individual clients. Revenue enhancements might involve increases in advisory fees that would be allowed under the Proposed Rule. At the present time, however, scant empirical information is available about (a) the proportion of current IRA-related revenues of broker-dealers that currently result from transactions that would be newly prohibited under the rule and (b) how much adoption of the rule would increase broker-dealer costs. As a result, we analyze potential effects qualitatively.

The two comments predict that if the rule were adopted, then retail IRA investors would be harmed in several ways. Specifically, the comments claim that these investors would be harmed by the following: less advice about which investment products to include in their IRA portfolios, higher minimum account balances, increased advisory fees, less help in setting up IRAs, decreased levels of contributions to their IRAs, and some investors closing their IRAs. Our analysis suggests that it is likely that some or all of these undesirable effects would result if the proportion of revenues currently derived from sources that would be newly prohibited by the rule, or if the cost increases attributable to the rule, would be sufficiently large.

We then consider how adoption of the rule could affect the well-being (or "utility") of IRA investors. We use a particular set of assumptions that seem plausible in light of economic theory and empirical information from a variety of studies. Although our specific conclusions are sensitive to the use of alternative sets of plausible assumptions, the general implications of our analysis appear to apply broadly. These implications include (a) effects of the Proposed Rule on IRA investors result from the effects of adopting the rule on several outcomes affecting investor welfare, such as fees, rates of return on IRA portfolios, amounts of time investors spend dealing with their IRAs, and levels of contributions that investors make to their IRAs; (b) effects of the Proposed Rule on investor well-being are likely to differ substantially across different types of IRA investors (such as those with different levels of financial capability); and (c) some investors may be made better off and others made worse off by adoption of the Proposed Rule. Available information, however, is inadequate to estimate how much any particular group of investors is likely to be helped or harmed by adoption of the rule.

Acknowledgments

This work was sponsored by the U.S. Department of Labor, Employee Benefits Security Administration. We thank our colleagues Noreen Clancy, Jack Clift, Jeffrey Garnett and Joanne Yoong for helpful comments and suggestions. The RAND Labor and Population review process employs anonymous peer reviewers, including at least one reviewer who is external to RAND. Two anonymous reviewers provided thorough and constructive feedback on the draft of this paper, and we are grateful to them for helping us improve it.

Abbreviations

CFED	Center for Financial and Economic Decision Making
CFR	Code of Federal Regulation
DOL	U.S. Department of Labor
EBSA	Employee Benefits Security Administration
ERISA	Employee Retirement Income Security Act
FK	Fischel and Kendall
IAR	Investment Advisor Representative
IRA	Individual Retirement Account
IRC	Internal Revenue Code
OW	Oliver Wyman
PTE	Prohibited Transaction Exemption
RIA	Registered Investment Advisor

Chapter 1. Introduction

In May 2011, 37 million (31.2 percent) and 18.6 million (15.7 percent) U.S. households owned traditional and Roth IRAs, respectively. The traditional IRA accounts had mean and median assets of $118,000 and $42,500, respectively; the corresponding figures for Roth IRAs were a mean of $41,100 and a median of $20,000 (Investment Company Institute, 2011a, pp. 2, 7). Encouraging U.S. households to save for retirement and helping them invest those funds wisely given their financial means and retirement goals are major social concerns. This paper considers these concerns in the context of a recent proposed policy initiative of the U.S. Department of Labor (DOL).

Specifically, in October 2010, DOL's Employee Benefits Security Administration (EBSA) proposed a rule that more broadly defined the circumstances under which organizations and individuals are considered to be a "fiduciaries" by reason of giving investment advice to an employee benefit plan, plan participant, or individual retirement account (IRA) holder (U.S. DOL, EBSA, 2010). In response to "The Definition of the Term 'Fiduciary'—Proposed Rule" (henceforth the Proposed Rule), the DOL received numerous comments, including two that predicted undesirable, unintended consequences for IRA investors if the rule is adopted. In this report, we review and assess the predictions contained in these two comments in response to a request from the EBSA.

Oliver Wyman (2011; henceforth OW) and Fischel and Kendall (2011; henceforth FK) raise several issues concerning how financial service providers may respond or react to the DOL's Proposed Rule. More specifically, the OW and FK comments focus on putative effects of the Proposed Rule in the markets for professional advice concerning traditional and Roth IRAs. Our analysis focuses on potential effects in these markets and the corresponding implications for investors in traditional and Roth IRAs.

The concerns raised by OW and FK pertain to two categories of potential effects of the Proposed Rule. First, OW and FK consider industry responses that would affect the prices, quality, and quantity of services provided to retail investors as they open IRAs, decide how much to contribute to them over time, and choose particular assets to include in their IRA portfolios. Second, OW and FK consider transition and ongoing costs borne in the first instance by the industry because of broker-dealer responses to adoption of the Proposed Rule.

The remainder of this paper is organized as follows. Chapter Two provides background on the Proposed Rule and other legal and regulatory issues. In Chapter Three, we then consider, and find unconvincing, FK's arguments that broker-dealers and their associated representatives rarely, if ever, self-deal despite financial incentives to do so because of competitive and reputational considerations. In Chapter Four, we describe our conceptual framework for considering potential industry responses to adoption of the Proposed Rule and employ this framework to review and assess several claims by OW and FK about industry responses. Chapter Five considers potential effects of adoption of the Proposed Rule on the well-being of retail investors, effects that are likely to differ substantially across investors. In Chapter Six, we offer an overview of our analysis and conclusions.

Chapter 2. Legal and Regulatory Background

By way of background for our analysis, this chapter describes the legal and regulatory context underlying the effects of the Proposed Rule on the opportunities and incentives of industry-side participants in the retail IRA markets, which we refer to as "broker-dealers and their associated representatives."[1] The Proposed Rule is intended to avoid problems that could potentially stem from conflicts of interest by prohibiting some forms of compensation that give these firms and individuals financial incentives to place their interests ahead of those of their clients.

The statutory foundation of the law pertaining to retail IRA markets is the Internal Revenue Code (IRC), which—subject to exceptions described presently—prohibits fiduciaries from engaging in transactions that involve self-dealing. A transaction is considered to involve self-dealing under the IRC if that transaction benefits the fiduciary financially. Under the IRC, violations of prohibitions on self-dealing are subject to an excise tax of 15 percent of the "amount involved"[2] or 100 percent of that amount if the violation is not corrected[3] in a timely fashion.[4]

The general prohibition of self-dealing, however, does have explicit exemptions; these are known as prohibited transactions exemptions (PTEs). For example, and perhaps most important for our analysis, PTE 86-128 allows fiduciaries to be compensated through "commissions" for transactions involving common investment products, such as mutual funds, securities, and insurance products. It is our understanding that for the purposes of PTE 86-128, allowed "commissions"[5] include investor payments that are

---

[1] ERISA defines a person as follows: "(9) The term 'person' means an individual, partnership, joint venture, corporation, mutual company, joint-stock company, trust, estate, unincorporated organization, association, or employee organization" (29 U.S.C. 1002). Thus, both natural and artificial persons can constitute fiduciaries.

[2] Section 4975(f)(4) of the IRC defines the term "amount involved," generally, as the greater of (a) the amount of money and the fair market value of the other property given or (b) the amount of money and the fair market value of the other property received in such transactions. For purposes of the first tier excise tax, the fair market value is determined as of the date on which the prohibited transaction occurs, whereas, for purposes of the second tier excise tax, the fair market value is the highest fair market value during the taxable period described in § 4975(f)(2).

[3] Section 4975(f)(5) of the IRC defines "correction" as undoing the transaction to the extent possible, but in any case placing the plan in a financial position not worse than that in which it would be if the disqualified person were acting under the highest fiduciary standards.

[4] The statute identifies the correction as being "timely" if it is done within the "taxable period," which in turn Section 4975(f)(2) defines as the period beginning with the date on which the prohibited transaction occurs and ending on the earliest of (a) the date of the mailing of a statutory notice of deficiency, (b) the date on which the first tier excise tax is assessed, or (c) the date on which correction of the prohibited transaction is completed.

[5] In the context of IRAs, the word "commission" means different things to different organizations such as the Employee Benefits Security Administration (EBSA), the U.S. Securities and Exchange Commission, and industry groups. Regarding industry groups, Investment Company Institute (2011a, p. 215, Glossary),

directly related to particular transactions (such as one-time front-end loads) but they do not allow ongoing payments (such as trailing commissions—periodic fees paid for investments held in an investor's IRA). Moreover, for the purposes of our analysis, we assume that implementation of the Proposed Rule would not result in the overturning or preemption of any existing PTEs.

Finally, although the actions that constitute prohibited transactions under IRC Section 4975 might also violate other laws that apply to fiduciaries (e.g., federal and state securities laws and state common law), being labeled a "fiduciary" under the Proposed Rule would not in itself provide a basis for meeting the definition of a fiduciary under these other laws. Personal liability on the basis of meeting the definition of a fiduciary under the Employee Retirement Income Security Act (ERISA) would come only from violation of the ERISA fiduciary standards with respect to ERISA plans[6] (but not IRAs). Similarly, liability based on the definition of "fiduciary" under the IRC would come only from violations of the fiduciary self-dealing and third-party compensation prohibitions in IRC 4975. Under the common law, which differs across states, fiduciary status typically confers on investment advisers and broker-dealers legal duties of loyalty and prudence (or "care") (U.S. Securities and Exchange Commission, 2011, pp. 45, 51), but these duties would arise from the broker-dealer meeting the definition of a fiduciary under the relevant state law, not under the proposed regulation. Nevertheless, we cannot rule out the possibility that this view will be challenged in court with at least some investment advisers and broker-dealers being sued and incurring litigation costs.

---

for example, defines "commission" as "[a] fee paid to a broker or other sales agent for services related to transactions in securities." To limit confusion in what follows, we usually avoid using the term "commission" in this review. A notable exception is that we do use the term "commission-based"—which is often used by OW and FK—to describe a common type of advisory relationship when using this term helps us explain claims by OW and FK and our assessments of those claims.

[6] Under ERISA, the two primary fiduciary duties can be described as follows: The duty of loyalty provides that a fiduciary shall discharge his or her duties with respect to a plan solely in the interest of the participants and beneficiaries and for the exclusive purpose of providing benefits to participants and their beneficiaries (ERISA section 404(a)(1)(A); codified at 29 U.S.C. 1104(a)(1)(A)). The duty of prudence provides that a fiduciary must act "with the care, skill, prudence and diligence under the circumstances then prevailing that a prudent man acting in a like capacity and familiar with such matters would use in the conduct of an enterprise of a like character and with like aims" (ERISA section 404(a)(1)(B); codified at 29 U.S.C. 1104(a)(1)(B)). In some circumstances, other fiduciary duties apply, such as a duty of candor or disclosure and a duty to follow the plan's requirements. The duty of prudence—which is violated for giving bad, rather than conflicted, advice is usually subjected to a judicial process inquiry. That is, a fiduciary generally cannot be sued for giving advice that turns out to be bad, unless the plaintiff can show that the advice was not the product of a reasonably prudent investigation. An often-quoted explanation is

> [C]ourts have construed the "prudent person standard" under ERISA as an "objective standard, requiring the fiduciary (1) to employ proper methods to investigate, evaluate and structure the investment; (2) to act in a manner as would others who have a capacity and familiarity with such matters; and (3) to exercise independent judgment when making investment decisions." (*United States v Mason Tenders Dist. Council of Greater N.Y.*, quoting *Lanka v O'Higgins*)

Chapter 3. Do Conflicts of Interest Currently Affect the Behavior of Financial Advisors?

The Proposed Rule is aimed at protecting investors from professional advice that fails to serve their interests because of the presence of financial rewards to broker-dealers and individuals providing professional advice for recommending and selling investments that do not best serve the objectives of their clients. FK (2011, p. 21, para. 49) makes explicit a helpful distinction in this regard, and this distinction is used as we go along. In particular, FK distinguishes "*potential* conflicts of interest" [emphasis added]—namely financial incentives for firms and advisors to deviate from advice that would best serve their clients' interests—from "*acting on* these conflicts to the detriment of customers" [emphasis added].

Neither FK nor OW dispute the existence of *potential* conflicts for "commissioned-based" broker-dealer models (advisory or discount). FK argues, however, that firms and advisors rarely, if ever, *act on* their conflicts of interest, summarizing this claim by writing: "market discipline protects investors" (FK, 2011, p. 22, para. 50). The FK argument is that market competition and financial benefits of firms and their representatives having reputations for faithfully serving their clients' interests prevent broker-dealers and advisors from acting on their conflicts in ways that harm their clients. The FK argument is unconvincing, however, as we explain presently.

Fischel and Kendall's Conceptual and Theoretical Arguments

FK (2011, p. 22, para. 51) argues that competition among broker-dealers enables those who put their clients' interests first (i.e., do not act on potential conflicts of interest) to gain customers from companies that do not similarly refrain from self-dealing. It seems, however, that this argument would pertain only to fairly sophisticated investors (or those taking cues from them). Less sophisticated investors are unlikely to be able to assess effectively which broker-dealers refrain from self-dealing or act on their conflicts of interests less often than their competitors do.

More specifically, FK (2011, p. 22, para. 51) writes that broker-dealers "can earn sales by pointing out" shortcomings of other firms and that there are substantial "opportunities for investors to readily seek 'second opinions.'" It is not clear, however, how credibly or at what cost broker-dealers can effectively "point out" the shortcomings of their competitors, which seems more difficult the less sophisticated investors are. Moreover, it is doubtful that investors who cannot assess the quality of the advice they receive[1] could identify and access sources of objective, unbiased, and credible second opinions or effectively determine what the costs of that advice would be.

FK (2011, pp. 24–25, para. 53) also discusses three theoretical articles[2] that illustrate the possibility of losing customers because of acting on conflicts of interest. Even FK's own descriptions of these studies, however, do not suggest that competition substantially reduces the prevalence of self-dealing or that it reduces its frequency to acceptable levels. More specifically, as described by FK, (a) Bolton, Freixas, et al. (2007) concludes that "competition reduces the gains from lying" (but not, we point out, that competition eliminates or even greatly reduces lying)[3]; (b) Krausz and Paroush (2002) concludes that competition makes it easier for "dissatisfied investors to transfer from one financial advisor to another" (but perhaps not, we point out, easy enough to impose major market penalties on those who act on their conflicts) and competition leading to business losses results in "reducing the incentive to deceive" (but may not, we point out, suffice to reduce deception substantially or to acceptable levels)[4]; and (c) Patron and Roskelley (2008) concludes that real estate "agents are less likely to suggest aggressive bargaining strategies [for their clients] when there is little market competition."

In sum, it seems likely that, other things equal, putting clients' interest first will tend to enable broker-dealers to gain new customers—albeit only to the extent that investors can accurately assess differences across broker-dealers in this regard. The theoretical arguments of FK, however, do not suggest that the strength of this incentive suffices to eliminate self-dealing or even to reduce its prevalence to acceptable levels.

---

[1] Australian Securities & Investment Commission (2006) found that Australian consumers were rarely able to detect "bad advice" induced by conflicts of interest involving advisor compensation, namely advice that was subjectively judged to have failed to consider important factual issues, did not fit the client's needs, or was likely to leave the client worse off.

[2] Strictly, the analyses in Krausz and Paroush (2002) are not entirely theoretical. In particular, in addition to offering and mathematically analyzing a theoretical model, to develop a better feel for the implications of their model, they perform simulations of the model calibrated with data from Israel.

[3] It is important to note how the reputation costs are modeled in Bolton, Freixas, et al. (2007). In particular, after selecting a product, an investor can immediately detect when he has been given bad advice, and a bank's reputation is known to all potential investors. In reality, it can be very difficult for investors to judge whether the advice they receive is good advice, especially for the large numbers of investors who struggle with financial literacy. Even if investors could make that judgment, it would most likely take several years to do so. The assumption that a bank's reputation could be known to all potential investors is likewise unrealistic, but evidence indicates that reputation and "word-of-mouth" are important in the financial service sector. For example, Hung, Clancy, et al. (2008) finds that more than 75 percent of investors who use a financial service provider found their provider through referral, either professional or personal.

[4] In the Krausz and Paroush (2002) model, highly risk-averse investors with relatively small accounts are particularly vulnerable to exploitation. The authors present simulation evidence calibrated to Israeli banking data that suggests increased competition may reduce the effects of conflicts of interest for investors who have high relative risk aversion, but it will not improve matters for investors with low relative risk aversion.

Empirical Evidence Offered by Fischel and Kendall to Support Their Claims

FK offers four pieces of empirical information to support their claim that competition greatly limits or eliminates acting on conflicts of interest. First, FK (2011, p. 22, para. 50) asserts that

> The revealed preference of investors themselves for commission-based investment services with respect to IRAs provides *perhaps the most powerful evidence* that, even in the presence of potential incentives for broker-dealers to behave opportunistically with respect to their clients, market discipline protects investors. [emphasis added]

Investors' choices of advisory services reflect, however—in addition to their preferences—such factors as fees, the sizes of their IRA portfolios, and the fact that many investors misunderstand how and how much they actually pay for advice. More specifically, large numbers of retail investors may currently choose "commission-based" advisory relationships—from among a limited range of alternatives, including entering into a fee-only advisory relationship and foregoing professional advice—largely because they fail to understand or appreciate the hidden costs of a commission-based relationship.[5] Moreover, many investors are unaware of how and how much they pay for financial advice ("Commissions Win the Day over Fees," 2011) and are confused about the nature of their advisory relationships (Hung, Clancy, et al., 2008).

Second, FK (2011, p. 22, para. 51, first bullet) argues that the industry is highly competitive because of the existence of thousands of broker-dealers competing to provide "financial planning and investment advice." The existence of thousands of firms seeking business from the same potential customers for the same or similar services does not, however, undermine or refute our view that competition among providers of investment advice may not suffice to eliminate acting on conflicts of interest. This is not sufficient because investor information about the quality of advice offered by different firms is far from perfect or complete (as is assumed in models of perfect competition, for example), which greatly limits the ability of the market to penalize firms that act on their conflicts of interest.

Third, FK (2011, p. 22, para. 51, second bullet) reports that "[p]ublic financial filings by companies in the industry consistently indicate a high degree of competition." This argument and the four specific examples of statements on financial firms' annual reports are unpersuasive. The word "competitive" means something different in the business world than it does in economic theory. More specifically, in the business world, the word "competition" connotes rivalry among firms seeking more business in a particular market, no matter how well the operation of this market serves consumers or promotes economic efficiency. In contrast, in economics the word "competition" often connotes

---

[5] Partially supporting our interpretation, namely that perceptions of the value of financial advice may differ substantially from actual value, FK (2011, p. 22, para. 50) writes, "[m]oreover, half of all mutual fund shareholders indicated they had ongoing relationships with financial advisers, illustrating the continuing value investors *perceive* in their relationships with financial service providers" [emphasis added].

"perfect competition," which is a set of idealized conditions (assumptions) under which markets produce economically efficient outcomes.[6]

Fourth, FK (2011, p. 22, para. 51, third bullet) reports that "[a]verage fees for mutual funds" have declined "dramatically and consistently" over time. Such trends in prices, however, are not in and of themselves evidence of extensive competition. This is because competitive models of pricing predict that (in equilibrium) prices will approximate costs, not that prices will decrease over time (unless costs decline over time).

Fischel and Kendall Arguments About Investor Benefits from Firms' *Reputational* Concerns

FK also considers the effects of reputational concerns on the prevalence of advisor self-dealing. Their arguments do not, however, provide any substantial reason to believe that reputational concerns—alone or in combination with competitive forces—result in acceptable levels of self-dealing.

Conceptual and Theoretical Arguments

FK (2011, p. 25, para. 54) writes:

> if a firm develops a reputation for low-quality service, its clients will be less likely to use that firm's services in the future, and will be less willing to recommend the firm's services to others. In the context of financial services firms, this provides an incentive for those firms to provide high-quality service to their clients, even in the presence of potential conflicts of interest.

We agree that firms with better reputations for providing "high-quality service to their clients" will, other things being equal, attract more customers and that this mechanism represents an incentive to increase the quality of their services, or at least those aspects of quality that are apparent to prospective customers. FK does not, however, claim that reputational concerns result in low or acceptable levels of self-dealing. In fact, as discussed above in the context of the effectiveness of competition in limiting self-dealing, the ability of broker-dealers to build reputations appears to be greatly limited by the inability of unsophisticated investors to assess the quality of the advice they receive.

Empirical Evidence Offered by Fischel and Kendall to Support Their Claims About Reputational Concerns

FK offers two examples of empirical evidence to support their claim that market rewards for having a good reputation with investors greatly limits or essentially eliminates acting on conflicts of interest.

[6] Moreover, statements to investors in SEC filings, such as form 10-Ks, about extensive competition seem to involve no substantial downside for reporting companies and could help protect them from investor and derivative lawsuits alleging that they misled investors in their financial filings.

First, FK (2010, p. 26, para. 56, first bullet) reports that many companies say in their U.S. Securities and Exchange Commission filings that their reputations are very important to them. It is likely to be true that broker-dealers do value good reputations with actual and potential clients. Even if true, however, this does not imply that many (or even any) companies have earned good reputations by consistently refraining from acting on their conflicts of interest.

Second, FK (2010, p. 26, para. 56, second bullet) reports that referrals from current or past clients are important to generate new business, and referrals will not occur if their reputations are not good. Again, even if referrals are highly desirable (as suggested by Hung, Clancy, et al., 2008) and are more common for broker-dealers that deserve good reputations, this does not imply that most, or even many, broker-dealers, actually do promote referrals by consistently refraining from acting on their conflicts of interest.

Implications of Other Literature

Findings of several studies that FK do not address in this context suggest that acting on conflicts of interest is much more commonplace than FK indicates.

First, the results of two studies that involved trained auditors posing as ordinary clients of financial advisors indicate that commissions and associated conflicts of interest lead to the provision of some advice that does not well serve clients' interests (Australian Securities & Investment Commission, 2006; Mullainathan, Noth, and Schoar, 2010). Moreover, Charles River Associates (2002) finds evidence of compensation mechanisms affecting market shares in directions that would be expected if advisors do act on their potential conflicts of interest.

Second, other studies indicate that the hidden costs of bad advice are reflected in lower rates of return for investors (Bergstresser, Chalmers, and Tufano, 2009; Christoffersen, Evans, and Musto, 2012). Note especially in this regard that even if—as those authors cannot rule out—investors receive intangible benefits from professional advice (e.g., time savings or peace of mind) along with lower returns, this does not contradict the evidence of lower returns.

Last, evidence from the psychology literature indicates that even when a person is attempting to behave ethically, he may not understand the extent to which self-interest is biasing his judgment (Moore and Loewenstein, 2004). Furthermore, Moore, Tanlu, and Braverman (2010) finds that once a biased opinion has been formed, providing self-interested incentives to be unbiased may not overcome the original bias, in part because individuals are unaware of the subtle effects bias have on their decisionmaking processes. Self-serving bias may also lead people to unconsciously overvalue evidence that supports their position and downplay evidence against (Bazerman, Morgan, et al., 1997).

Summary

FK argues that acting on potential conflicts of interest is too rare to raise substantial policy concerns, and, thus, that potential benefits to investors of implementing the

Proposed Rule would be insubstantial. Their arguments and the purported evidence supporting them is unconvincing, and findings from studies FK do not discuss support the view that acting on potential conflicts of interest is fairly widespread and is costly to investors. Thus, it is not appropriate to dismiss the Proposed Rule as obviously failing a cost-benefit test on the basis that there could be at most minor investor benefits to adopting it.

Chapter 4. Industry Responses to Adoption of the Proposed Rule

In this chapter, we first develop a conceptual framework to consider how the industry may respond to the adoption of the Proposed Rule. We then use this framework to assess OW and FK claims about the effects of adopting the Proposed Rule. We also use this framework in Chapter Five to analyze potential effects of the rule on the well-being of retail IRA investors.

We define the effects of the Proposed Rule on any outcome as the level of that outcome *with* the rule (equivalently, if the Proposed Rule is adopted) minus the level of that outcome *without* the rule (equivalently, if the Proposed Rule is not adopted). Thus, we—along with the Department of Labor in the Proposed Rule, OW, and FK—conceptualize for cost-benefit analysis purposes comparing the state of the world under the Proposed Rule with the *status quo*.[1]

Adoption of the Proposed Rule would eliminate some sources of revenue currently available to nonfiduciaries and would increase broker-dealer costs of serving the retail IRA market. The likely effects of such changes on the service offerings and fees facing retail investors can be analyzed fairly straightforwardly using widely accepted concepts in microeconomics (specifically, the theories of firms and markets).

To analyze industry responses to implementation of the Proposed Rule, we assume the following:

1. Current rules, competition, and investors' imperfect information about which broker-dealers offer their clients the best values result in a status quo industry equilibrium in which the total revenues of broker-dealers associated with services provided to retail IRA investors at least cover the total costs of providing these services and, because of the information problems facing investors, may involve total revenues exceeding their costs (because competition is not sufficiently effective to eliminate all economic profits).
2. If the Proposed Rule is adopted, some of the IRA-related total revenues of broker-dealers would be lost to them, namely, the revenues associated with transactions involving (organizational and individual) advisor-compensation arrangements that would be newly prohibited under the rule.
3. If the Proposed Rule is adopted, broker-dealers will incur some additional costs associated with compliance and protection from lawsuits alleging failure to satisfy their legal duties associated with fiduciary status.

[1] Analyses of the effects of the Proposed Rule on outcomes are examples of *positive* (or "descriptive") economic analysis, which is often described as "analysis of what *will* be." In contrast, *normative* (or "prescriptive") analyses are often described as "analysis of what *should* be." A cost-benefit analysis is a normative analysis that assumes that the social goal is improving economic efficiency. Predictions about how the policy under consideration will affect outcomes of social concern—that is, positive analyses—are critical components of a cost-benefit analysis.

Broker-dealers will not serve the retail IRA market unless they expect to be able to earn sufficient revenues to cover their costs of doing so.

An implication of these assumptions is that if broker-dealers do not respond to the implementation of the Proposed Rule by changing service offerings to reduce costs or increase revenues from allowed sources, they may not be able to cover their costs. For a broker-dealer who would earn positive profits under the status quo, the revenue losses and cost increases attributable to the rule might not be sufficiently large for the firm's total revenues to fall short of total costs, in which case no response to the Proposed Rule may be required for such a firm to continue to offer the services it currently offers for the same fees. Even such broker-dealers, however, may have to change their service offerings and fee schedules to continue to be willing to serve the retail IRA market because of decreases in their profits result from their competitors' responses to the implementation of the rule. In sum, to continue to offer advisory services to retail IRA investors—equivalently, to cover costs and remain in this business—broker-dealers may have to reduce costs (which in large measure might require reducing levels of service[2]), increase revenues by setting higher levels of fees that are not prohibited, or both.[3]

Industry Responses to Increased Costs and Elimination of Some Revenue Sources

For purposes of discussion, we define two terms pertaining to distinct characteristics of advisory services. First, the "quality" of advisory services pertains to the extent to which broker-dealers and their representatives refrain from self-dealing. More specifically, we will say that the "quality" of advice is higher the greater is the extent to which advisors put their clients' interests first when providing advice. Second, the "intensity" of advisory services pertains to the time and effort devoted by advisors to advising each IRA client on average, with intensity increasing with increases in per-client time and effort.

In discussing potential industry responses to adoption of the Proposed Rule, OW and FK limit their consideration of responses by implicitly assuming that no fundamentally new business models will emerge to serve the retail IRA market as a result of adopting the rule. More specifically, their comments implicitly assume that under the rule the only kinds of advisory relationships will be (see OW, 2011, pp. 5–7) the kinds of relationships that currently exist: (a) fee-based advisory relationships and (b) full-service and discount ("commission-based") brokerage relationships. Within that assumed realm, the comments consider only how broker-dealers may respond by changing (in our terminology) the quality and intensity of the services provided and direct fees paid by investors. In essence, then, OW and FK are imagining that under the rule, broker-dealers will search for sets of services (at different levels of intensity) they could offer to retail IRA investors

[2] The pressures to reduce cost if the Proposed Rule is implemented might lead to greater efforts to operate more efficiently—that is, to provide the same levels of service at lower cost.

[3] Note, however, that increasing fees will reduce revenues if demand is elastic.

along with associated levels of allowed fees that they would expect to cover costs given what their competitors are expected to offer and how investors' demand is likely to respond.[4]

Within the conceptual framework just sketched, changes from the status quo of the industry equilibrium as a result of adoption of the Proposed Rule are the combined effects of two outcomes of new fiduciary status: (a) higher costs for broker-dealers and (b) reductions in their revenues. We begin by considering potential effects on costs.

Additional Industry Costs Resulting from New Fiduciary Status

FK considers several kinds of cost increases that could result from the adoption of the Proposed Rule. First, FK argues that the industry or its associated advisors might incur substantial costs of additional training and certification required for advisors to continue to give advice if the Proposed Rule were adopted. More specifically, FK (2011, p. 6, para. 13) writes

> many representatives of broker-dealer firms that currently provide services to IRA investors do not currently hold the certifications necessary to operate as fee-based investment advisors, and that if the Proposed Rule were implemented, these representatives would need to gain additional certification in order to continue to serve their clients or attract new clients.

In this context, FK (2011, p. 6. para.14) refers to the "Series 65" license. It does appear that—to be allowed to charge fees for providing investment advice if the Proposed Rule were adopted—many financial advisors who would become fiduciaries would need to become certified as investment advisor representatives (IARs) or registered investment advisors (RIAs). The requirements for IAR or RIA certification differ, however, from state to state.[5] Passing the Series 65 exam (or the Uniform Investment Advisor Law

---

[4] The possible emergence of fundamentally new business models for serving IRA investors would be especially salient, if as OW and FK predict (see Major Contractions in the Investment Advice Industry below), adoption of the Proposed Rule would result in major declines in the size of the industry. The potential financial rewards for developing and implementing viable new business models are likely to be extremely large; thus, we anticipate that both incumbents and potential new providers of advice to retail IRA investors would devote substantial time and effort to trying to develop new, profitable business models. This effort would be challenging even for industry insiders with unusually rich knowledge of the industry, costs, behavior of retail investors, and so on. Thus, we can do no more than speculate about what new business models might emerge. One seemingly plausible possibility is that incumbents or new entrants will offer in unbundled form services that are currently offered only as packages (or "bundled") by broker-dealers. Unbundling might, for example, involve separate provision of help in setting up IRAs, deciding how much to save, choosing investment products, and so on. Such unbundling would allow investors to purchase only the services they think are worth the costs to them, which could promote economic efficiency. Perhaps, however, some unbundled services cannot be priced in a way that will find enough customers to cover the costs of providing these services. If so, organizations that are not motivated by profit if they participate in this market (e.g., employers, foundations, other not-for-profits, government agencies) may offer such services for no charge or for prices that many investors would be willing to pay, thereby mitigating any harmful effects of industry responses for at least some retail IRA investors.

[5] See, generally, North American Securities Administrators Association, *Exam FAQs*, n.d.

exam)—along with passing a background check and paying fees—suffice for certification in most states. Most of those states, however, allow individuals to qualify without passing the Series 65 exam if they have passed other specified exams or hold other specific certifications.[6]

FK (2011, p. 6, para. 14) reports that an average of 50 hours per applicant is a "conservative estimate" of the preparation time required to take the Series 65 exam. We have searched but found no studies of the costs of obtaining Series 65 certification. RIA Compliance Consultants[7] (*Frequently Asked Questions About the Series 65 Examination*, n.d.), however, advises people to "[p]lan to spend between 45 and 60 hours and at least four weeks of studying for the series 65 examination." This admittedly limited evidence suggests that the figure of 50 hours offered by FK is reasonable.[8]

FK (2010, p. 4, para. 11) also writes, "[u]ndoubtedly, financial service providers and their representatives would incur significant compliance costs in complying with new regulations." Additional compliance costs would include, for example, costs associated with broker-dealers developing new corporate policies, communicating these policies to

---

[6] See, North American Securities Administrators Association, "When Can I Register as an Investment Adviser Representative if I Haven't Taken the Series 65, or Series 66 in Combination with the Series 7?" *Exam FAQs*, n.d.

> Most states will allow an individual to substitute one of the following certifications for passing the exam: CFP – Certified Financial Planner (granted by the CFP Board of Standards); CIC – Chartered Investment Counselor (granted by the Investment Adviser Association); ChFC – Chartered Financial Consultant (granted by the American College); PFS – Personal Financial Specialist (granted by the American Institute of Certified Public Accountants); and CFA – Chartered Financial Analyst (granted by the Chartered Financial Analyst Institute).

[7] RIA Compliance Consultants describes itself as follows:

RIA Compliance Consultants is a team of industry experienced professionals dedicated to working with investment advisors who are also committed to implementing good compliance and risk management strategies. By working together, RIA Compliance Consultants helps investment advisors navigate the maze of investment advisor compliance regulations and find the best ways to satisfy their obligations. (RIA Compliance Consultants, *About Us*, n.d.)

[8] The 50 hours of preparation time per examination taker assumption is one piece of FK's illustrative calculation of the industrywide cost of more than $295 million (which they characterize as "necessarily a rough approximation") derived by assuming that (a) 60 percent of 300,000 registered representatives nationwide would take the Series 65 exam once each, and (b) the value of an hour of preparation time is the "2009 median hourly wage of personal financial advisers" of $32.79 (FK, 2011, pp. 6–7, para. 14). Regarding (a), we have no basis for second-guessing the 300,000 figure (for which no source is provided by FK). We should expect that the number of people choosing to take the Series 65 exam will not be dramatically different from the number of people who can be employed giving financial advice if the Proposed Rule is adopted. We have no basis, however, for assessing the degree to which numbers of professional financial advisors would decline in response to adoption of the Proposed Rule. Regarding (b), using a wage rate (or average hourly earnings) of an individual to approximate the social cost of an hour of that person's time is standard practice in cost-benefit analysis, and using the median hourly wage rather than the mean seems appropriate (and, perhaps, a bit conservative) in that the median is not sensitive to very high figures that are likely to characterize personal financial advisors who do not serve the retail IRA market. Finally, we have no basis for judging the representativeness of Primerica's ratio of series 65 license holders (namely, 233) to registered representatives (16,000); note, however, that this ratio is not used in the calculation of the $295 million estimate.

their associated advisors, and monitoring compliance. Moreover, FK (2011, p. 5, para. 11) suggests that firms will incur additional costs of defending lawsuits enabled by their fiduciary status and "fiduciary liability insurance." As discussed in the context of legal and regulatory background, broker-dealers may well incur additional litigation costs because of their newly conferred fiduciary status. Moreover, broker-dealers may purchase additional liability insurance in response to what they perceive as additional liability exposure resulting from adoption of the rule.

Lost Revenues Resulting from Prohibition of Some Revenue Sources

If the Proposed Rule were adopted, (compliant) broker-dealers would lose some proportion of the IRA-related revenues they currently receive. Neither OW nor FK provide any information about the revenue sources that would be prohibited if the Proposed Rule were adopted or their importance to broker-dealers; in fact, these questions are currently unanswerable. Nonetheless, it appears OW and FK assume that if the Proposed Rule were adopted, broker-dealers would lose a substantial proportion of all of their IRA-related "commission-based" revenues.

Effects of the Proposed Rule on Investors Claimed by Oliver Wyman and Fischel and Kendall

OW and FK predict numerous adverse impacts on investors from the adoption of the Proposed Rule, including increased fees, increased minimum account balances, and reduced access to advice. In the following sections, we discuss each of these possibilities.

Increases in Fees

OW and FK state in various places in their comments that investors' "expenses" or "costs" associated with receiving advice would increase as a result of the Proposed Rule.[9] It seems clear from their comments, however, that the investors' "expenses" or "costs" to which they refer are limited to fees paid directly to broker-dealers by retail IRA investors. Stated differently, the "direct fees" on which OW and FK focus exclude other investor costs of receiving advice from broker-dealer representatives, such as indirect fees.[10] Moreover, if retail investors obtain lower returns (or bear excessive risk) on their investments because of self-dealing, then these shortfalls in returns are also properly

---

[9] For example, FK title a section of their comments "Expenses Paid By IRA Investors Would Likely Rise Significantly Due to the Proposed Rule" (FK, 2011, p. 7).

[10] OW acknowledges the existence of indirect fees in writing that

> the current brokerage model that has developed to serve IRA accounts is incompatible with this requirement [under the Proposed Rule that "firms and their associated representatives may not receive different levels of compensation based on investment choices of retail investors in protected IRA accounts"], often involving both direct and indirect fees, such as shareholder servicing fees, sales and distribution fees, revenue-sharing and other fees. (OW, 2011, p. 14)

viewed as investor costs of receiving advice. In sum, the *full costs* to investors of a "commission-based" brokerage arrangement may currently comprise direct fees, indirect fees (some of which may be prohibited if the rule is adopted), and reduced investment performance (because of self-dealing). Thus, an assessment of the effects of adopting the rule on investors should consider the effects on indirect fees and investment performance along with effects on direct fees. In considering investor costs or expenses, however, OW and FK focus on direct fees.[11]

Both OW and FK claim that direct fees (although they use the terms "costs" and "expenses") paid by IRA investors should be expected to increase as a result of adopting the Proposed Rule. This claim is plausible as long as the rule would eliminate a substantial portion of broker-dealers' current revenue or substantially increase their costs. This is true because (according to our conceptual framework) elimination of some revenue sources or increased compliance and litigation costs would be expected to lead broker-dealers to seek out ways to increase revenues, and direct fees appear to be one of the leading candidates for consideration in that regard.

Moreover, in several places in their comments, OW and FK (a) argue that adoption of the Proposed Rule would lead to large increases in investors' costs or expenses associated with receiving advice through broker-dealers, and (b) conclude that these cost or expense increases will price many retail investors out of the market for IRA-related financial advice. It appears that the latter argument (b) may be correct despite the fact that the validity of the former (a) is questionable if "costs" and "expenses" include indirect as well as direct fees.

It seems likely that substantial increases in direct fees—which may or may not result from adoption of the rule—could lead to substantially reduced demand for professional financial advice by retail investors no matter what happens to full costs as a result of the adoption of the Proposed Rule. Demand could be reduced because direct fees appear to be the most (and perhaps for many investors, the only) salient component of full costs; for such investors, direct fees are the key "price" to which their demands for advisory services respond.[12] More specifically, many investors are likely to *perceive* large price

---

[11] For example, while considering only direct fees, FK (2011, pp. 11–12, para. 24) provides a numerical example purporting to demonstrate high costs to investors resulting from the adoption of the Proposed Rule. Specifically, FK (2011, p. 11, para. 24) assumes that 66 percent of IRAs "are held in accounts with commission-based brokerages" and claims that

> if the Proposed Rule led to even a 1 basis point increase in annual costs relative to assets for these investors, it would generate $277 million (= 4.2 trillion X 66% X 0.01%) in additional expenses for investors annually, or over $2 billion over 10 years in current dollars, using a discount rate of 7%.

This example appears incomplete, however, in that it ignores potential decreases in *indirect* fees and improvements in portfolio performance that might also result from adoption of the Proposed Rule.

[12] Barber, Odean, et al. (2005) argues that over the past several decades, investors have become less willing to invest in funds with high front-end load fees, but they have not become more sensitive to operating expenses despite a dramatic increase in the latter.

increases (namely increases in direct fees) no matter what the effect of the rule is on their full costs of receiving professional financial advice.[13]

Increases in Minimum Account Balances

FK also predicts that adoption of the Proposed Rule will lead to increases in minimum account balances, thus preventing some IRA holders from continuing to receive professional advice. Specifically, FK (2011, p. 12, para. 25) writes, "[b]ecause the revenue generated by low balance accounts is small, an increase in costs would likely mean that these firms would increase minimum account balance requirements for IRA investors." We have no empirically grounded basis for predicting whether minimum account balances would increase if the Proposed Rule were implemented. Within our conceptual framework, however, minimum account balances would increase if (and, perhaps, only if) broker-dealers would be unable to create a set of service offerings that could be purchased by IRA investors at fee levels that could cover broker-dealers' costs, while retaining the minimum account balances that exist under the status quo.

Eliminating "Commission-Based" Advisory Services

OW predict that full-service or discount brokerage arrangements, which they also refer to as "commission-based," will cease to exist, and IRA investors will be offered only fee-based advice. Specifically, OW (2011, p. 13) writes: "*Under the proposed rule, Oliver Wyman expects retail brokerage firms would presume that current brokerage account and service offerings would create a fiduciary duty, and respond by limiting the provision of help and investment services to the fee-based advisory model only*" [italics in original].

Thus, OW predicts that if the Proposed Rule were adopted, then IRA investors would not be able to receive professional investment advice through what they call "commission-based" advisory arrangements and would be left with fee-based advice as their only alternative. If so, then another claim of OW and FK would imply that many or most IRA investors would be harmed by adoption of the rule. Specifically, in different places in their reports OW and FK state that retail investors "prefer" commission-based arrangements with their financial advisors, thereby suggesting that investors who would not have access to such arrangements if the Proposed Rule were implemented would necessarily be worse off.

In support of their claim about investor preference, OW (2011, p. 11) writes, "[i]nvestors represented in the study group overwhelmingly use the brokerage

---

[13] FK (2011, pp. 13–14, para. 29), citing Ernst & Young (2011, p. 7), writes, "[a]s reported by analysts at Ernst & Young, under the new rules . . . it seems likely that the mass market and the typical bank customer will not be enthusiastic about paying the sort of fees that make offering the advice attractive." This claim is consistent with our analysis—despite the fact that the Ernst & Young (2011) report pertains to policy developments in the United Kingdom, not in the United States—as long as we interpret "will not be enthusiastic about paying" as "will decline to pay and thus forego services" and if adoption of the Proposed Rule would, in fact, lead to substantial increases in direct fees.

relationship model as opposed to a fee-based advisory model, with 22.4 million or 88% holding brokerage IRAs." As discussed, however, choices made by investors reflect not only their preferences but also their opportunities and available information. In the present context, this means that most IRA investors have chosen "commission-based" arrangements because of their preferences as well as other considerations, including the fees they believe they are paying for advice and the sizes of their IRA portfolios.

Major Contractions in the Investment Advice Industry

FK (2011, pp. 13–14, para. 29) writes

> As reported by analysts at Ernst & Young, under the new rules . . . simplified advice becomes a major economic challenge, requiring a radically reduced cost base if it is to present a solution for the mass market. . . . There is a real possibility that the independent advisory sector, as we know it, will shrink significantly. (citing Ernst & Young, 2011, p. 9)

We think it is likely that many broker-dealers and investment advisors would exit the IRA market if the Proposed Rule were adopted only if the proportion of current revenues that would have to be replaced or increases in costs were fairly substantial.

The number of professional advisors needed to serve the IRA market would be expected to decrease as a result of adopting the rule to the extent that broker-dealers exit the IRA market or take other steps to reduce their IRA-related advisory activities. The size of any effect of implementing the Proposed Rule on numbers of professional advisors serving the IRA market cannot be predicted, however, because of a lack of sufficient empirical information. Even 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.

Other Predicted Reductions in Service Offerings

FK and OW suggest that reductions in service offerings, higher direct costs for IRA investors, higher minimum account balances, and decreases in the numbers of investment advisors will result in IRA investors receiving fewer services of different types. More specifically, they predict that retail investors will

- Receive less help in setting up IRAs (because providing such help is costly to broker-dealers, and fewer IRA investors will pay the prices required to receive advice about IRAs)
- Receive less investment advice (because fewer retail investors will have professional investment advisors)

- Reduce their IRA contributions and open fewer new IRAs (FK, 2011, p. 12, para. 25, writes, "because, as described above, prices for financial services would likely rise, the Proposed Rule would be likely to cause some individuals to choose not to open IRA accounts or to invest less in them.")
- Close some IRAs (OW, 2011, p. 16, writes, "we believe that a large number of small balance investors forced to switch to a new firm to access the advisory channel would be likely to take a cash distribution rather than successfully re-invest in a new IRA.").

We cannot assess these claims for two major reasons. First, we have no basis for gauging the extent of the adjustments that broker-dealers would have to make—which may include cost reductions, revenue increases, or both—to operate profitably if the Proposed Rule were adopted. Second, even if we did know how much broker-dealers would have to reduce costs, increase revenues, or both to operate profitably under the rule, we would have no basis for predicting whether the cost reductions would involve lesser availability of particular advisory services that currently are bundled in "commission-based" advisory relationships, such as those claimed in the four bullet points.

Chapter 5. How Would Adoption of the Proposed Rule Affect the Well-Being of Retail IRA Investors?

In this chapter, we consider how the adoption of the Proposed Rule would affect the well-being (or "utility") of retail IRA investors. To do so, we draw on discussions about potential industry responses and several studies discussed in Burke, Hung, et al. (forthcoming). We conclude that adoption of the rule may make some investors better off (increase their utility) and other investors worse off (decrease their utility), but available empirical information is inadequate to quantify these effects for any particular group.

As we detail in this chapter, available information—both theoretical and empirical—suggests that the Proposed Rule would differentially affect IRA investors, depending on their levels of financial capability. Thus, it seems necessary to consider separately the welfare effects on individuals with high and low financial capability. In principle, to quantify effects of the Proposed Rule on investors, one would want to (a) identify groups of investors for which the welfare effects are likely to be similar within groups and dissimilar across groups, (b) estimate in dollar terms the average change in utility within each group, and (c) determine the number of investors in each group. With that information, the social benefits or costs of the rule associated with effects on retail IRA investors could be estimated.[1] Available empirical information, however, is inadequate for quantifying the welfare effects on particular investors, no less quantifying the value of the aggregate gains to those investors who would benefit and the aggregate losses to those who would be made worse off.

An Analytic Approach to Drawing Inferences About Effects on Retail IRA Investors

To analyze welfare effects for different groups of investors, we decompose the question of overall welfare effects on investors into pieces we can think about (theorize) directly and, ideally, about which existing literature provides useful guidance.

In this analysis, we focus on the effects of the adoption of the Proposed Rule on retail IRA investors who would receive professional financial advice without the rule. We further assume that without adoption of the rule, (a) some advisors will self-deal to some extent and (b) investors who do receive advice will follow it and make some investments that are not best suited to their circumstances and retirement goals. Moreover, we assume

[1] More specifically, the social benefits or costs of adoption of the rule on all retail IRA investors jointly could then be calculated as follows. First, estimate for each investor group the number of investors in the group and the (monetized) average effect of adoption of the rule on the utility of investors in that group. Second, for each group, multiply the number of investors by the average effect on those investors, which would represent the total social benefits or costs for that group. Finally, add up over groups the welfare effects for each group.

that all relevant investors who receive professional advice do so through what OW (2011, p. 6) calls a "brokerage" relationship (full service or discounted).

To simplify the analysis and exposition, we explicitly consider only four outcomes underlying the well-being of a retail IRA investor:

1. Actual total (direct plus indirect) fees paid to receive professional advice, whether or not these total fees are understood by the investor
2. Rate of return on the investor's IRA portfolio
3. Time spent making financial decisions regarding IRAs
4. Amounts contributed to IRAs.[2]

Thus, the first outcome (total fees paid) and any reductions in the second outcome (rate of return) resulting from advisor self-dealing account for the "full costs" of receiving advice as defined above. The third outcome (time spent) involves investor time devoted to making decisions about IRAs, some of which may be saved by receiving professional advice. The fourth outcome (amounts contributed) exemplifies several other outcomes whose levels could be sensitive to whether an investor receives professional advice, including "coaching" to contribute amounts consistent with the investor's age, retirement goals, and so on.[3]

An Illustration of the Approach

To analyze the issue at hand, one must (explicitly or implicitly) make additional assumptions. We now detail the assumptions we use to analyze effects on investor well-being. This set of assumptions appears plausible to us on the basis of our review of pertinent theoretical and empirical literature. Undoubtedly, other sets of assumptions also would be plausible. Thus, our analysis illustrates a general approach to thinking about effects of the adoption of the rule on investor welfare. At the least, using this particular set of assumptions enables us to illustrate the complexity of the question of aggregate welfare effects and the kinds of factors that are expected to determine this outcome. We present and discuss the illustration in sufficient detail to enable readers to alter our assumptions and consider the qualitative implications of alternative assumptions.

In what follows, we assume that

1. For the same service levels of advice as would exist without the rule, with the rule, direct fees would increase (to make up broker-dealer revenues lost on newly prohibited transactions and any cost increases).
2. For the same service levels of advice as would exist without the rule, with the rule, indirect fees would decrease (namely indirect fees associated with transactions that would be newly prohibited if the rule were adopted).

---

[2] We are not aware of any empirical evidence that financial advisors actually do influence investors' decisions about how much to contribute to their IRAs. We nonetheless consider the possibility that advisors do, in fact, induce many investors to contribute more than they otherwise would contribute because industry sources claim that this is true, and we think it is plausible.

[3] Other potential desirable effects of coaching include opening up IRAs and not closing them.

3. Because of the first assumption, and the apparent focus of many investors on direct fees (because indirect fees are not understood by or are not salient to many investors[4]), if the rule were adopted, some investors (who would receive advice without the rule) will choose to forego professional advice,[5] thus decreasing the *quantity* of (or the number of IRA investors receiving) professional advice.
4. The *quality* of professional advice would increase if the rule were adopted because of less self-dealing by advisors, implying higher rates of return for those who continue to receive advice than they would experience without the rule.[6]
5. The *intensity*[7] of professional advice would
   a. *Decrease* for those continuing to receive advice (e.g., less coaching to contribute more to IRAs) as a result of broker-dealers choosing to lower their costs of providing advice
   b. *Decrease to zero* under the rule for those foregoing advice.
6. On balance, total advisory fees borne by investors receiving advice would decrease because of reductions in indirect fees and advice intensity (outweighing increases in direct fees).
7. The two types of IRA investors are "sophisticated" and "unsophisticated" investors:
   a. *Sophisticated* investors who receive advice follow that advice, which lowers their rates of return on their IRAs when their advisors self-deal. They do not use (or benefit from) professional advice (coaching) to open IRAs, to keep open existing IRAs, or to contribute amounts appropriate to their

[4] Barber, Odean, et al. (2005) empirically investigates investor responses to different kinds of fees and argues that over time investors have become increasingly sensitive to front-end loads (which are often large and salient), but investors have not become sensitive to operating expenses (which can be masked by the volatility of fund returns). Many investors are confused about the total cost of financial advice ("Commissions Win the Day over Fees," 2011) and are confused about the nature of their advisory relationship (Hung, Clancy, et al., 2008).

[5] The literature provides some guidance regarding factors other than fees that influence whether investors choose to receive professional advice. Hung and Yoong (2010) finds that people with low financial ability are more likely to solicit advice about investment decisions, and Hackethal, Inderst, and Meyer (2010) finds that investors are more likely to rely on advice when they are less confident in their own financial expertise. In addition, other research suggests that the market for advice may be somewhat segmented, with many capable individuals investing autonomously and with many less capable investors (Gil-Bazo and Ruiz-Verdú, 2008, 2009) or investors who receive extra utility from advice (Del Guerico, Reuter, et al., 2010) relying on financial advisors. An exception is the working paper by Calcagno and Monticone (2011), which finds that higher financial literacy reduces the probability of investing autonomously. Finally, Hortacsu and Syverson (2004) finds that investors who purchase through brokers may be less able or less willing to invest on their own.

[6] Substantial empirical evidence suggests that financial advisors are influenced by their compensation schemes (Australian Securities & Investment Commission, 2006; Mullainathan, Noth, and Schoar, 2010) and that investors who purchase through advisors earn lower returns than those who invest autonomously. See, for example, Bergstresser, Chalmers, and Tufano (2009) and Christoffersen, Evans, and Musto (2012).

[7] As defined in Chapter Four, "intensity" refers to the extent of advice received by each investor (intensive margin), not the number of investors receiving advice, which we refer to as "quantity" of advice (extensive margin). In particular, we have defined "intensity" as the average amount of time and effort per IRA client that advisors spend working on behalf of these clients.

circumstances and retirement goals. In sum, with or without the Proposed Rule, sophisticated investors benefit from professional advice *only* in terms of saving time,[8] and without the rule, they sacrifice investment returns because their advisors act on their conflicts of interest.

b. As is assumed about sophisticated investors, *unsophisticated* investors benefit from time savings (with or without the rule), and without the rule, they lose potentially higher returns on their investments because of advisor self-dealing. Unlike sophisticated investors, however, without the rule, unsophisticated investors also benefit from help in choosing investment products (relative to making these choices without professional advice) and from coaching to open IRAs, contribute appropriate amounts, and so on.

Thus, we have defined *sophisticated* and *unsophisticated* investors implicitly in terms of which of the activities of professional advisors help or harm them.[9] Our assumptions imply that without the rule *sophisticated* investors would earn returns at least as high as those they receive through a financial advisor if they were to invest on their own.[10]

Table 5.1 presents our theory- and literature-based conclusions about likely (qualitative) effects of the adoption of the Proposed Rule (again, outcome with the rule minus outcome without the rule). It considers (in different rows) four groups of IRA investors, namely the four combinations defined by whether investors would continue to receive professional advice with the rule (receive or forego) paired with the two types of investors (sophisticated or unsophisticated). For each of these four situations, we consider (across the columns) *effects of adopting the rule* on the four outcomes that underlie the utilities of investors: total advisory fees, rates of return on an IRA portfolio, time spent by investors dealing with their IRAs, and the amounts investors contribute to their IRAs. Each cell in the table reports our (necessarily tentative) conclusion about how an outcome would be affected by the adoption of the rule.

To further explain the table and illustrate the kind of reasoning involved in filling out its cells, consider first the case of a sophisticated investor who would continue to receive advice if the rule were adopted. As indicated in the table, we conclude that if the rule were adopted, such an investor would benefit—other things equal—from lower total advisory fees than without the rule. This reduction in advisory fees would result from reductions in service intensity because of the efforts of broker-dealers to reduce the cost of serving each retail IRA investor. Moving across the columns, we also predict that this investor (a) would benefit in terms of rate of return on his or her IRA portfolio because of

[8] Although low financial capability appears to positively associated with reliance on advice, some investors who use a financial advisor may be capable but time constrained (Investment Company Institute, 2007; Hackethal, Haliassos, and Jappelli, 2011).

[9] Whether a particular investor chooses to forego professional advice and how well he or she does is likely to be affected by the availability of many resources available to all investors such as the online "retirement savings calculator" (Kiplinger, Retirement *Savings Calculator,* n.d.) and other resources that can be found, for example, by a Google search of "how much do I need to save for retirement?"

[10] Of course, sophistication would more realistically be assumed to take on many discrete values or to vary along a continuum (i.e., take on many, not just two, values). The assumption of only two types of investors is a useful distortion of reality that simplifies the discussion while allowing us to make our key points.

reduced prevalence of advisor self-dealing (implying higher quality advice) because of adoption of the rule, (b) might increase time spent dealing with his or her IRA to make up for broker-dealers reducing the intensity of their services, and (c) would experience no change in IRA contributions because (by assumption) sophisticated investors do not change their behavior in response to coaching. Putting the pieces together, whether a sophisticated investor who continues to receive advice is made better off or worse off by adoption of the rule depends on the combined effects of the utility changes associated with the effects on the four outcomes. In this case (first row of the table), one would conclude that the investor is better off under the rule as long as the decrease in total advisory fees and the increase in rate of return because of the adoption of the rule increase utility by more than the decrease in utility associated with any extra time the investor would spend dealing with his or her IRA.

Next consider a sophisticated investor who would forego professional advice if the rule were adopted. As indicated in (the second row of) the table, total advisory fees would decrease to zero because no professional advice would be received, and (by assumption) it would not have an effect on IRA contributions. Moreover, we think it likely that this investor would benefit from a higher rate of return on his or her IRA portfolio by not having to rely on conflicted advice but would tend to be worse off because of having to spend more time dealing with his or her IRA.[11] In sum, this type of investor would be better off if the rule were adopted only if the sum of the utility increments associated with lower fees and higher rates of return would outweigh the decrement in utility due to spending more time dealing with his or her IRA.

Turning to unsophisticated investors, first consider those investors who continue to receive professional advice under the rule. As indicated in the third row of Table 5.1, we would expect (as with sophisticated investors who continue to receive advice) that (a) their total advisory fees would be lower if the rule were adopted, (b) the rate of return on their portfolios would be higher because of less self-dealing than without the rule, and (c) time spent might increase somewhat. Unlike sophisticated investors (who are assumed not to respond to coaching), these unsophisticated investors might decrease their levels of contributions to their IRAs because of reduced advisor time and effort devoted to coaching. In sum, this type of investor would benefit from the rule only if the increments to utility from lower fees and higher returns would outweigh the utility decreases associated with extra time spent dealing with their IRAs and decreased contributions to their IRAs.

The fourth and final case is that of unsophisticated investors who forego professional advice. These investors would benefit from reduction of advisory fees to zero. It is unclear whether the rates of return on their portfolios would be higher or lower than under the status quo if the rule were adopted, however. More specifically, although unsophisticated investors likely experience lower returns than are possible without the

[11] Hackethal, Haliassos, and Jappelli (2011) suggests one reason investors may be willing to receive lower returns through a financial advisor is that they are too busy to manage their investments on their own.

rule because of advisors acting on conflicts of interest, unsophisticated investors may or may not be capable of getting higher returns than they would without the rule if they forego professional advice. Finally, these investors would tend to be made worse off by adoption of the rule because of (a) additional time spent dealing with IRAs (if they keep them open or, looking further into the future, open IRAs at all) and (b) lower levels of contributions to their IRAs (assuming, as we do, that investors tend to contribute less than the amounts that would maximize their utility).

Summary

This chapter has analyzed qualitatively how adoption of the Proposed Rule would affect the well-being of different types of retail IRA investors who receive advice without the rule (i.e., under the status quo). To develop and illustrate our main points, we employed a particular set of assumptions that seem plausible based on economic theory and available empirical evidence. The main points are that (a) effects of the Proposed Rule are likely to differ substantially across different types of retail IRA investors, (b) effects of adoption of the rule on investor well-being depend on effects of the rule on several outcomes, and (c) available information is inadequate to conclude whether any particular kind of investor would be made better off or worse off by adoption of the rule.

We have argued that welfare effects for particular investors depend on the effects of adopting the Proposed Rule on several outcomes. In our illustrative application, these outcomes are total advisory fees, rates of return on IRA portfolios, time spent by investors dealing with IRAs, and—as an example of potential effects of coaching by advisors—levels of contributions to IRAs. We have argued that the effects of adopting the Proposed Rule would differ depending on investor sophistication and, if the rule were adopted, whether investors would choose to continue to receive professional advice. For all four of the situations analyzed, we found that adopting the rule not only would tend to improve at least one outcome for an investor but also would worsen at least one of the other four outcomes. Thus, it seems that any aggregate prediction of the likely effects of the rule on the welfare of retail IRA investors would best consider multiple types of investors for whom adoption of the Proposed Rule would affect multiple determinants of investor well-being. We are unable, however, to conclude whether any particular kind of investor would be made better or worse off by adoption of the rule.

Chapter 6. Conclusion

This paper reviews and qualitatively evaluates the economic predictions contained in comments from OW and FK in response to the DOL's "Definition of the Term 'Fiduciary'—Proposed Rule." We also propose and illustrate a conceptual approach to analyzing effects of adoption of the Proposed Rule on investor welfare.

One of the comments argued at length that—because of broker-dealers' incentives stemming from competition and reputational concerns—financial advisors rarely, if ever, act on their conflicts of interest. If so, there would be little, if any, benefit to adoption of the rule. We find this claim and supporting arguments and evidence unpersuasive. Moreover, substantial empirical evidence indicates that many financial advisors are influenced by their compensation schemes in ways that harm investors.

The comments also predicted that adoption of the Proposed Rule would substantially increase the costs borne by retail IRA investors, particularly because of increases in direct fees. Direct fees for advice might well increase in response to adoption of the rule, although we have no empirical basis for assessing the likelihood of such an effect. Because—as indicated by other empirical literature—direct fees are highly salient to investors, a substantial increase in direct fees would likely result in fewer investors choosing to receive professional financial advice. Even if direct fees do increase substantially, however, this does not mean that investors' full costs of advice would necessarily increase. Indirect fees could decline and IRA portfolio performance could improve because of rule-induced reductions in advisor self-dealing.

Additionally, investors who continue to work with a financial advisor may receive less advice if broker-dealers respond to adoption of the rule by reducing the amount of time and effort devoted to each client to reduce their costs of providing IRA-related advice. Investors who no longer receive financial advice are likely to have to spend more time managing their retirement savings.

A key question is how adoption of the rule would affect the well-being or welfare of investors. Effects of adopting the rule on IRA investors could be different for different types of investors and also could depend on whether investors would receive professional advice if the rule were adopted. Previous research and a simple framework that we propose and illustrate suggest that the effects of the Proposed Rule depend on the effects of adoption of the rule on several outcomes affecting the well-being of IRA investors. Such outcomes include total fees, rates of return on IRA portfolios, time investors spend dealing with their IRAs, and the amounts investors contribute to their IRAs. It may be that adopting the rule would help some IRA investors and harm others. Without considerably more information, we cannot predict whether the benefits to IRA investors would outweigh their costs for any particular type of IRA investor or in the aggregate.

Table 5.1. Effects of Implementation of the Proposed Rule on Sophisticated and Unsophisticated Investors (effects = levels with the rule minus level without rule)

A. Sophisticated Investors

	Component of Investor Well-Being			
	Total Advisory Fees[a]	**Rate of Return on IRA Portfolio**	**Time Spent by Investor**	**IRA Contributions (Due to Professional Coaching)**
Continue to receive advice	Lower (due to intensity decreases) but still positive	Higher (due to less self-dealing than without the rule)	No difference or higher (higher would be due to decreased intensity)	No change
Forego advice	Lowered to zero	Higher (due to absence of self-dealing when no advice is received)	Higher (no help from advisors)	No change

B. Unsophisticated Investors

	Component of Investor Well-Being			
	Total Advisory Fees	**Return on Portfolio**	**Time Spent by Investor**	**IRA Contributions (Due to Professional Coaching)**
Continue to receive advice	Lower (due to intensity decreases) but still positive	Higher (due to less self-dealing than without the rule)	No difference or higher (higher would be due to decreased intensity)	Lower (due to intensity decrease)
Forego advice	Lowered to zero	Change could be positive or negative[b]	Higher (no help from advisors)	Larger decrease than for those receiving advice (coaching is effective for these investors)

SOURCE: Authors' Analyses.

[a] These are actual fees paid whether or not they are understood by investor; that is, we are assuming that it is actual fees that affect welfare even if it is only direct fees that affect investor choices regarding whether to pay to receive professional advice (because of lack of understanding of indirect or hidden fees).

[b] It is far from clear whether unsophisticated investors would earn higher returns on their own than they would if they receive conflicted advice.

References

Australian Securities & Investment Commission, *Shadow Shopping Survey on Superannuation Advice*, Report No. 69, 2006. As of May 30, 2012: http://www.asic.gov.au/asic/pdflib.nsf/LookupByFileName/shadow_shop_report_2006.pdf/$file/shadow_shop_report_2006.pdf

Bazerman, M. H., K. P. Morgan, et al., "The Impossibility of Auditor Independence," *Sloan Management Review*, Vol. 38, No. 4, 1997, pp. 89–94.

Barber, B., T. Odean, et al., "Out of Sight, Out of Mind: The Effects of Expenses on Mutual Fund Flows," *Journal of Business*, Vol. 78, No. 6, 2005, pp. 2095–2119.

Bergstresser, D., J. M. R. Chalmers, and P. Tufano, "Assessing the Costs and Benefits of Brokers in the Mutual Fund Industry," *Review of Financial Studies*, Vol. 22, No, 10, 2009, pp. 4129–4156.

Bolton, P., X. Freixas, et al., "Conflicts of Interest, Information Provision, and Competition in the Financial Services Industry," *Journal of Financial Economics*, Vol. 85, No. 2, 2007, pp. 297–330.

Burke, J., A. A. Hung, J. W. Clift, S. Garber, and J. K. Yoong, *Impacts of Conflicts of Interest in the Financial Services Industry*, Santa Monica, Calif. RAND Corporation, forthcoming.

Calcagno R., and C. Monticone, *Financial Literacy and the Demand for Financial Advice*, Moncalieri, Italy: Center for Research of Pensions and Welfare Policies, CeRP Working Paper, 2011. As of May 30, 2012: http://dx.doi.org/10.2139/ssrn.1884813

Charles River Associates, *Polarisation: Research into the Effect of Commission Based Remuneration on Advice*, 2002. As of November 1, 2011: http://www.fsa.gov.uk/pubs/other/pol_res1.pdf

"Commissions Win the Day over Fees," *American Banker*, Vol. 176, No. 94, 2011, p. 8.

Christoffersen, S. K., R. B. Evans, and D. K. Musto, *What Do Consumers' Fund Flows Maximize? Evidence from Their Brokers' Incentives*, Charlottesville, Va: Darden School of Business Working Paper No. 1393289, March 8, 2012.

Del Guercio D., J. Reuter, et al., *Broker Incentives and Mutual Fund Market Segmentation*, Cambridge, Mass.: National Bureau of Economic Research, NBER Working Paper Series, No. 16312, 2010.

Employee Retirement Income Security Act of 1974 (ERISA) (Pub. L. No. 93-406, codified in part at 29 USC § 1002 et seq.)

Ernst & Young, *RaDaR—Life and Pensions Outlook for 2011*, London, UK: Ernst & Young, 2011.

Fischel, D. R., and T. D. Kendall, "Comment to the Department of Labor on a Proposed Rule Regarding Fiduciary Status under ERISA," *Compass Lexecon*, April 12, 2011.

Gil-Bazo, J., and P. Ruiz-Verdú, "When Cheaper Is Better: Fee Determination in the Market for Equity Mutual Funds," *Journal of Economic Behavior & Organization*, Vol. 67, No. 3–4, 2008, pp. 871–885.

Gil-Bazo, J., and P. Ruiz-Verdú, "The Relation Between Price and Performance in the Mutual Fund Industry," *The Journal of Finance*, Vol. 64, No. 5, 2009, pp. 2153–2183.

Hackethal, A., M. Haliassos, and T. Jappelli, *Financial Advisors: A Case of Babysitters?* Frankfurt, Germany: Goethe University Frankfurt, 2011.

Hackethal, A., R. Inderst, and S. Meyer, *Trading on Advice*, Washington, D.C.: Center for Economic and Policy Research, CEPR Discussion Paper: 8091, 2010.

Hortacsu, A., and C. Syverson, "Product Differentiation, Search Costs, and Competition in the Mutual Fund Industry," *Quarterly Journal of Economics*, Vol. 119, No. 2, 2004, pp. 403–456.

Hung, A., N. Clancy, J. Dominitz, E. Talley, C. Berrebi, and F. Suvankulov *Investor and Industry Perspectives on Investment Advisers and Broker-Dealers*, Santa Monica, Calif.: RAND Corporation, TR-556-SEC, 2008. As of June 1, 2012:
http://www.rand.org/pubs/technical_reports/TR556.html

Hung, A., and J. Yoong, *Asking for Help: Survey and Experimental Evidence on Financial Advice and Behavior Change*, Santa Monica, Calif.: RAND Corporation, WR-714-1, 2009. As of June 1, 2012:
http://www.rand.org/pubs/working_papers/WR714-1/

Internal Revenue Code § 4975 (codified at 26 USC § 4975 - Tax on prohibited transactions).

Investment Company Institute, "Why Do Mutual Fund Investors Use Professional Financial Advisers?" *Research Fundamentals*, Vol. 16, No. 1, 2007. As of May 30, 2012:
http://www.ici.org/pdf/fm-v16n1.pdf

Investment Company Institute, *2011 Investment Company Fact Book*, Washington, D.C.: Investment Company Institute, 2011a.

Investment Company Institute, *ICI Research Perspective*, Vol. 17, No. 8A, November 2011b.

Kiplinger, *Retirement Savings Calculator*, n.d. As of July 26, 2012:
http://www.kiplinger.com/tools/retirement-savings-calculator.html

Krausz, M., and J. Paroush, "Financial Advising in the Presence of Conflict of Interests," *Journal of Economics & Business*, Vol. 54, No. 1, 2002, pp. 55-71.

Lanka v O'Higgins, 810 F.Supp. 379, 387 (N.D.N.Y. 1992).

Moore, D. A., and G. Loewenstein, "Self-Interest, Automaticity, and the Psychology of Conflict of Interest," *Social Justice Research*, Vol. 17, No. 2, 2004, pp. 189-202.

Moore, D. A., L. Tanlu, and M. H. Braverman, "Conflict of Interest and the Intrusion of Bias," *Judgment and Decision Making*, Vol. 5, No. 1, 2010, pp. 37–53.

Mullainathan, S., M. Noth, and A. Schoar, *The Market for Financial Advice: An Audit Study*, Cambridge, Mass.: National Bureau of Economic Research, Working Paper Series, No. 17929, 2010.

North American Securities Administrators Association, "When Can I Register as an Investment Adviser Representative if I Haven't Taken the Series 65, or Series 66 in Combination with the Series 7?" *Exam FAQs*, n.d. As of August 8, 2012: http://www.nasaa.org/industry-resources/exams/exam-faqs/#13

Oliver Wyman, *Assessment of the Impact of the Department of Labor's Proposed "Fiduciary" Definition Rule on IRA Consumers*, New York: Oliver Wyman Group, April 12, 2011.

Patron, H. E., and K. S. Roskelley, "The Effect of Reputation and Competition on the Advice of Real Estate Agents," *Journal of Real Estate Finance and Economics*, Vol. 37, No. 4, 2008, pp. 387–399.

RIA Compliance Consultants, *About Us,* n.d. As of July 26, 2012: http://www.ria-compliance-consultants.com/ria_compliance_consultants.html

RIA Compliance Consultants, *Frequently Asked Questions About the Series 65 Examination*, n.d. As of July 26, 2012: http://www.ria-compliance-consultants.com/investment_advisor_investment_adviser_rep_series_65_exam_faq.html

United States Department of Labor, Employee Benefits Security Administration, "Definition of the Term 'Fiduciary'—Proposed Rule," *Federal Register*, Vol. 75, No. 204, October 22, 2010, pp. 65263–65278.

United States Securities and Exchange Commission, *Study on Investment Advisers and Broker-Dealers*, Washington, D.C.: U.S. Securities and Exchange Commission, January 2011.

United States v Mason Tenders Dist. Council of Greater N.Y., 909 F.Supp. 882, 886 (S.D.N.Y. 1995).

49

THE JOURNAL OF FINANCE • VOL. LXIV, NO. 5 • OCTOBER 2009

The Relation between Price and Performance in the Mutual Fund Industry

JAVIER GIL-BAZO and PABLO RUIZ-VERDÚ*

ABSTRACT

Gruber (1996) drew attention to the puzzle that investors buy actively managed equity mutual funds, even though on average such funds underperform index funds. We uncover another puzzling fact about the market for equity mutual funds: Funds with worse before-fee performance charge higher fees. This negative relation between fees and performance is robust and can be explained as the outcome of strategic fee-setting by mutual funds in the presence of investors with different degrees of sensitivity to performance. We also find some evidence that better fund governance may bring fees more in line with performance.

MANY STUDIES HAVE attempted to determine whether equity mutual funds are able to consistently earn positive risk-adjusted returns.[1] Although these studies have documented significant differences in risk-adjusted returns across funds, it became apparent early on (Sharpe (1966)) that those differences are to a large extent attributable to differences in fund fees. Most research on mutual funds has thus been aimed at determining whether the cross-sectional variation in performance that is not explained by fees can be explained by the existence of managers with superior stock-picking skills (see, e.g., Chevalier and Ellison (1999)). However, little attention has been paid to the relation between before-fee performance and fees. In this paper, we focus on this relation and investigate whether differences in fees reflect differences in the value that mutual funds create for investors.

*Gil-Bazo and Ruiz-Verdú are with the Department of Business Administration, Universidad Carlos III de Madrid. The authors thank Campbell Harvey (the editor), an anonymous associate editor, and an anonymous referee for their comments, which have led to substantial improvements in the paper. We are also grateful to Laurent Barras, Clemens Sialm, Massimo Guidolin, Alfonso Novales, Mikel Tapia, Manuel F. Bagüés, Frédéric Warzynski, and Ralph Koijen as well as seminar participants at Tilburg University, IE Business School, University of the Basque Country, Universidad de Navarra, Universitat Pompeu Fabra, Universidad Carlos III de Madrid, Financial Management Association European Conference, and the annual meeting of the European Financial Management Association for helpful comments and suggestions. The usual disclaimer applies. Part of the paper was written while Javier Gil-Bazo was visiting the Finance Department of Tilburg University. The financial support of Spain's Ministry of Education and Science (SEJ2005-06655, SEJ2007-67448, and Consolider-Ingenio CSD2006-16), the BBVA Foundation, and the Madrid Autonomous Region (CCG07-UC3M/HUM-3413) is gratefully acknowledged.

[1] See, for example, Brown and Goetzmann (1995); Gruber (1996); Carhart (1997); Daniel et al. (1997); Wermers (2000); Cohen, Coval, and Pastor (2005); Kacperczyk, Sialm, and Zheng (2005); or Kosowski et al. (2006).

The Relation between Price and Performance in the Mutual Fund Industry
Author(s): Javier Gil-Bazo and Pablo Ruiz-Verdú
Source: *The Journal of Finance*, Vol. 64, No. 5 (Oct., 2009), pp. 2153-2183
Published by: Wiley for the American Finance Association
Stable URL: http://www.jstor.org/stable/27735168

Mutual fund fees pay for the services provided to investors by the fund. Because the main service provided by a mutual fund is portfolio management, fees should reflect funds' risk-adjusted performance. It follows that there should be a positive relation between before-fee risk-adjusted expected returns and fees. In contrast to this prediction, we find a puzzling negative relation between before-fee risk-adjusted performance and fees in a sample of U.S. equity mutual funds: Funds with worse before-fee risk-adjusted performance charge higher fees. To check the robustness of this finding, we use different estimation methods and performance measures, and we investigate the relation for different subsamples. The negative relation between before-fee risk-adjusted performance and fees survives these robustness checks.

We then set out to explain this anomalous relation by investigating the role of performance in the determination of fund fees. Christoffersen and Musto (2002) propose that mutual funds' fees are set taking into account the elasticity of the demand for their shares, so that funds facing less elastic demand charge higher fees. These authors argue that funds with worse past performance face less elastic demand because performance-sensitive investors leave funds following bad performance. If performance is persistent for at least the worse-performing funds (as indicated by Carhart (1997)), Christoffersen and Musto's hypothesis could explain our finding of a negative relation between fees and before-fee performance. Gil-Bazo and Ruiz-Verdú (2008) provide a related explanation. These authors set forth a model of the market for mutual funds in which competition among high-performance funds for the money of sophisticated (performance-sensitive) investors pushes their fees down and drives the low-performance funds out of that segment of the market. Low-performance funds then target unsophisticated investors, to whom they are able to charge higher fees. We also consider the possibility that funds with low expected performance have higher fees because they incur higher marketing costs, which they pass on to investors. Underperforming funds will have high distribution costs or advertising outlays if they target unsophisticated investors, and these investors are more responsive to advertising or more likely to use brokers to purchase mutual funds. We test these strategic fee-setting hypotheses against an alternative cost-based explanation, according to which the negative relation between performance and fees that results from univariate regressions would simply be due to the omission of fund characteristics associated with both lower operating costs and better performance.

To test these hypotheses, we estimate the relation between fund fees and performance, investors' sensitivity to fund performance, which we measure by the estimated slope of the relation between performance and money flows, and a number of variables that have been previously identified as determinants of funds' operating costs. Our results are consistent with the strategic explanations described earlier: Even after controlling for a host of fund characteristics, underperforming funds and funds faced with less performance-sensitive investors charge higher marketing and nonmarketing fees.

The apparent inability of mutual fund competition to ensure an adequate relation between risk-adjusted performance and fees raises the question of whether improvements in mutual fund governance can bring fees more into line with the value that funds generate for investors. To evaluate the potential effects of recent regulatory reforms that impose stricter requirements on mutual fund governance, we analyze the role played by fund governance in fee determination. We find some evidence that funds with boards of directors expected, a priori, to provide more effective protection of investors' interests charge fees that better reflect their risk-adjusted performance.

The paper is organized as follows. In Section I, we describe the mutual fund fee structure and the data set. In Section II, we explain how we estimate fund performance. In Section III, we estimate the relation between before-fee performance and fees and perform several tests to evaluate the robustness of the results. In Section IV, we discuss and test several explanations for the estimated relation between fees and performance. Section V concludes.

I. Data

A. Mutual Fund Fee Structure

Fund management fees are typically computed as a fixed percentage of the value of assets under management.[2] These fees, together with other operating costs, such as custodian, administration, accounting, registration, and transfer agent fees—comprise the fund's expenses, which are deducted on a daily basis from the fund's net assets by the managing company. Expenses are usually expressed as a percentage of assets under management known as the "expense ratio." Fees paid to brokers in the course of the fund's trading activity are detracted from the fund's assets, but are not included in the expense ratio.

Funds often charge "loads," which are one-time fees that are used to pay distributors. These loads are paid at the time of purchasing ("front-end load") or redeeming ("back-end load" or "deferred sales charge") fund shares and are computed as a fraction of the amount invested.[3]

Since 1980, funds may charge so-called "12b-1 fees," which are included in the expense ratio and, like loads, are used to pay for marketing and distribution costs. Since the 1990s, many funds offer multiple share classes with different combinations of loads and 12b-1 fees. Among the most common classes are class A shares, which are characterized by high front-end loads and low annual 12b-1 fees, and class B and C shares, which typically have no or low front-end loads but have higher 12b-1 fees and a contingent deferred sales load. This contingent deferred sales load decreases the longer the shares are held and is eventually eliminated (typically after 1 year for class C shares, and after 6 to 7 years for class B shares).

[2] Some funds allow the percentage to depend on fund performance. Although our data do not allow us to identify these funds, the evidence reported in Elton, Gruber, and Blake (2003), Kuhnen (2005), and Warner and Wu (2006) suggests that for most of our sample period the fraction of funds with incentive fees is very small.

[3] Funds often waive at least a fraction of the loads. Therefore, the loads typically reported in databases, such as the one we use in this paper, can often overestimate effective loads.

B. Sample Description

We obtain our data from the Center for Research in Security Prices (CRSP) Survivor-Bias Free U.S. Mutual Fund Database for the period from December 1961 to December 2005 (see Carhart (1997), Elton, Gruber, and Blake (2001), and Carhart et al. (2002) for detailed discussions of the data set). The initial sample contains all open-end mutual funds that are active in the 1961 to 2005 period. From this initial sample, we exclude all funds that we cannot confidently describe as diversified domestic equity mutual funds. Thus, we remove money market, bond and income, and specialty mutual funds, such as sector or international funds. To obtain our sample of diversified domestic equity mutual funds, we use the information on funds' investment objectives available in the CRSP database. Unfortunately, this information is not consistent throughout the 1961 to 2005 period. To create a homogeneous sample for the full sample period, we combine all the information available on funds' investment objectives. Some of our results, however, are derived only for the 1992 to 2005 period, for which the information on funds' investment objectives is precise and consistent.

We remove from the sample observations with no information on returns or expenses, or with zero expenses. The remaining sample contains some observations with extreme values for expenses or returns that are either data errors, or that correspond to small funds with unusually high expenses or very high volatility. Given the large size of the data set, we use Hadi's (1994) outlier detection method to search for these outliers and remove them from the sample.

Finally, to ensure that our results are not driven by differences between index and actively managed funds or between institutional and retail funds, we identify passively managed (index) and institutional funds and exclude them from the sample. In the Internet Appendix,[4] we provide a more detailed account of the procedure we use to construct our final sample, as well as summary statistics of the main variables.

C. Fund Governance Data

We use the January 2007 Morningstar Principia CD to obtain data on mutual funds' board quality. This data set includes several governance ratings as of December 2006, which Morningstar uses to compute the so-called "Stewardship Grade" for mutual funds. The measure of board quality that we use in our analysis (Morningstar's "board quality" grade) is the sum of four equally weighted components that measure, respectively, the degree to which the board has taken action "in cases where the fund clearly has not served investors well"; the significance of independent directors' investments in the fund; whether the board is "overseeing so many funds that it may compromise the ability to diligently protect the interests of shareholders"; and whether the fund meets the Securities and Exchange Commission (SEC) requirement for the proportion of

[4] The Internet Appendix is available at http://www.afajof.org/supplements.asp.

Table I

Summary Statistics: Board Quality Measure

The table shows the distribution of the board quality grade provided by Morningstar for funds in the governance subsample in year 2005. The first row reports the frequency of each grade; the second row reports the relative frequency (in percentage terms); and the third row reports the cumulative frequency (in percentage terms).

	Very Poor	Poor	Fair	Good	Excellent	Total
Frequency	1	69	374	486	176	1,106
%	0.09	6.24	33.82	43.94	15.91	100
Cumulative	0.09	6.33	40.14	84.09	100	

independent directors, regardless of whether it is subject to the requirement (see Morningstar (2006), pp. 1–2, for a detailed description). Thus, the board quality grade is closely related to board characteristics that have been the focus of both regulatory reform and academic research.[5]

Of the 3,677 actively managed, diversified, noninstitutional funds (fund classes) that are active in 2005 in our sample, there is board quality information in Morningstar's January 2007 Principia CD for only 1,106 funds (the governance subsample). Although only one-third of the funds in the sample belong to the governance subsample, these funds manage almost 80% of the total net assets managed by all funds in the sample. We note that the governance subsample is not a random sample from the whole population. Among other differences, funds in the governance subsample perform significantly better on average, belong to larger management companies, and are cheaper, larger, and older than those in the nongovernance subsample.

Table I reports the distribution of board quality grades. There are five grades to which Morningstar assigns a numerical score that ranges from zero to two: Very Poor (0), Poor (0.5), Fair (1), Good (1.5), and Excellent (2). As Table I shows, most funds have Fair or Good grades, and only one obtains a Very Poor grade. The resulting average grade lies between Fair and Good.

II. Mutual Fund Performance Estimation

We use Carhart's (1997) four-factor model to estimate before-fee risk-adjusted performance:

$$r_{it} = \alpha_i + \beta_{rm,i} rm_t + \beta_{smb,i} smb_t + \beta_{hml,i} hml_t + \beta_{pr1y,i} pr1y_t + \varepsilon_{it}, \qquad (1)$$

where r_{it} is fund i's before-expense return in month t in excess of the 30-day risk-free interest rate—proxied by Ibbotson's 1-month Treasury bill rate;[6] rm_t is the

[5] Morningstar's Stewardship Grade includes other components. See Wellman and Zhou (2008) for a recent analysis of the Morningstar Stewardship Grade.

[6] Because fund returns are reported after expenses, to retrieve monthly before-expense returns, we add annual expenses divided by 12 to reported returns. This measure is only an approximation

market portfolio return in excess of the risk-free rate; and smb_t and hml_t denote the return on portfolios that proxy for common risk factors associated with size and book-to-market, respectively. The term $pr1y_t$ is the return difference between stocks with high and low returns in the previous year. We include this term to account for passive momentum strategies by mutual funds.[7] The term α_i is the fund's *alpha* and captures the fund's before-fee risk-adjusted performance. We also consider Fama and French's (1993) three-factor model, which uses only rm_t, smb_t, and hml_t, as well as conditional versions of the four-factor model.

As in Carhart (1997), we follow a two-stage estimation procedure to obtain a panel of monthly fund risk-adjusted performance estimates. In the first stage, for every month t in years 1967 to 2005, we regress funds' before-fee excess returns on the risk factors over the previous 5 years. If less than 5 years of previous data are available for a specific fund-month, we require the fund to be in the sample for at least 48 months in the previous 5 years, and then run the regression with the available data. In the second stage, we estimate a fund's risk-adjusted performance in month t as the difference between the fund's before-expense excess return and the realized risk premium, defined as the vector of betas times the vector of factor realizations in month t.[8]

Rolling regressions yield a total of 232,386 monthly risk-adjusted before-fee returns corresponding to 3,109 different actively managed retail funds over 468 months. Although the average annualized monthly return before expenses in our sample equals 10.52%, subtracting the risk-free rate and the part of fund returns explained by the portfolio's exposure to the Fama–French three factors yields an average annualized monthly alpha of −21 basis points (bp), which is further reduced to −70.6 bp when we take momentum into account. The corresponding annualized standard deviations are 18.13%, 7.33%, and 7.15%, respectively.

III. The Relation between Fees and Performance

In a well-functioning mutual fund market, mutual fund fees should be positively correlated with expected before-fee risk-adjusted returns. Further, in the absence of market frictions, all funds should earn zero expected after-fee risk-adjusted returns in equilibrium since, otherwise, there would be excess demand (supply) for funds with positive (negative) expected after-fee risk-adjusted returns (Berk and Green (2004)). In this context, if investors know funds' alphas,

because we ignore the compounding effect of the accrual of expenses over the year, and because the actual accrual of expenses may not be completely smooth (Tufano, Quinn, and Taliaferro (2006)).

[7] Data are downloaded from Kenneth French's web site, http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/.

[8] We follow Fama and MacBeth (1973) in our choice of a 5-year estimation period, instead of the 3-year period used by Carhart (1997). Although a longer estimation period excludes a greater fraction of funds from the sample, it also reduces sampling error in betas and mitigates the effect of two forms of selection bias that affect mostly the subset of young funds in the CRSP database: omission bias and incubation bias (Elton, Gruber, and Blake (2001) and Evans (2009)).

then equilibrium requires that $\alpha_i - f_i = 0$ for every fund i, where f_i denotes fund i's fees, expressed as a fraction of the fund's assets. This equilibrium condition can be equivalently written as $\alpha_i = f_i$, for every fund i. Therefore, a graph depicting the equilibrium relation between fees and before-fee performance should yield an increasing linear relation with a slope of one. If, on the other hand, investors do not know funds' alphas, then α_i in the equilibrium condition is replaced by investors' expectation of fund i's risk-adjusted returns.

Equilibrium in the mutual fund market can be achieved through fee adjustment if funds with higher expected before-fee risk-adjusted returns increase their fees or underperforming funds lower theirs. However, Berk and Green (2004) show that in the market for mutual funds, market clearing can also be achieved via quantity adjustment: If there are decreasing returns to scale in fund management, then flows of money into funds that are expected to perform better will reduce those funds' expected performance until expected after-fee risk-adjusted returns are equalized across all funds in equilibrium. Whether market clearing takes place via fees, quantities, or a combination of both is, however, not material for the definition of market equilibrium. In any case, equilibrium requires that expected after-fee risk-adjusted returns be zero for all funds, and thus implies the linear relation (with a slope of one) between fees and before-fee performance described earlier.

To investigate the relation between fund fees and before-fee risk-adjusted performance, we first estimate by pooled ordinary least squares (OLS) the regression equation

$$\hat{\alpha}_{it} = \delta_{0t} + \delta_1 f_{it} + \xi_{it}, \quad i = 1, \ldots, N, \quad t = 1, \ldots, T, \tag{2}$$

where f_{it} is the fund's expense ratio and $\hat{\alpha}_{it}$ is its risk-adjusted before-fee performance measured according to Carhart's (1997) model.

The first row of Table II reports the slope coefficient and White's (1980) heteroskedasticity-robust standard error estimated using the whole sample of diversified actively managed retail equity funds. The regression includes month dummies to ensure that the estimated slope coefficient captures the cross-sectional relation between fees and risk-adjusted returns, not the effect of potentially correlated trends in those variables. The estimated slope coefficient is -0.63 and we can reject the null hypothesis of a unit slope at any conventional significance level. Thus, estimation of equation (2) yields results that are in stark contrast with the implications of a frictionless competitive market for equity mutual funds.

In a market with frictions, it is not clear whether a priori we should expect δ_1 to be greater or smaller than one. In one plausible scenario, better funds charge higher fees, but those fees are not high enough to fully compensate for the differences in before-fee performance. In this scenario, funds with higher fees offer a higher after-fee performance and the estimated δ_1 is greater than one ($\delta_1 > 1$ implies that increases in fees are matched by larger increases in before-fee performance). In another plausible scenario, better funds overcharge for their ability to generate returns, which leads to differences in fees that exceed

Table II

Before-Fee Risk-Adjusted Performance and Expense Ratios

The table shows estimated slope coefficients for the OLS regression of funds' monthly before-fee risk-adjusted performance on monthly expense ratios in the period from January 1962 to December 2005. Betas are estimated using Carhart's four-factor model (rows 1–3) or Fama–French's three-factor model (row 4) with a 5-year estimation period. Risk-adjusted performance in month t is estimated as the difference between the fund's monthly before-expense return in month t and the product of betas and the factor realizations for that month. All regressions include dummies for months. Standard errors are reported in parentheses and adjusted R^2 statistics in percentage. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Superscripts a, b, and c denote that the null hypothesis of a unit coefficient is rejected at the 10%, 5%, and 1% significance levels, respectively. The number of observations is 232,386.

Risk-adjusted Performance	Standard Errors	Coefficient	Adj. R^2
Carhart	White	$-0.6284^{***,c}$ (0.1055)	10.07
Carhart	Clustered by Time	$-0.6284^{**,c}$ (0.2529)	10.07
Carhart	Fama–MacBeth	$-1.4077^{***,c}$ (0.3352)	0.05
Fama–French	Clustered by Time	-0.2076^{c} (0.2599)	9.43

differences in performance and to an estimated $\delta_1 \in (0, 1)$. In this context, funds with higher fees should exhibit better before-fee performance but worse after-fee performance. Finally, fees could be completely unrelated to funds' before-fee performance, leading to $\delta_1 = 0$. However, the estimated slope coefficient is negative and significantly different from zero, which suggests an a priori much less plausible scenario in which funds with worse before-fee performance charge higher fees.

To account for cross-sectional correlation of residuals, we follow Petersen (2009) and Thompson (2006) and compute robust standard errors clustered by month. The second row of Table II shows that the robust standard error clustered by month (0.25) is more than twice as large as the White standard error (0.11), which suggests the presence of cross-sectional correlation in residuals (Petersen (2009)). However, further clustering by both fund and month (to also account for serially correlated residuals) barely changes the standard error (0.27). Therefore, unless otherwise noted, throughout the rest of this section we report robust standard errors clustered by time in all pooled OLS regressions. We also estimate the relation between fees and performance using the Fama–MacBeth two-step approach (Fama and MacBeth (1973)), which is designed to correct for cross-sectional correlation of residuals: First, we estimate monthly regressions, and then we use the resulting monthly slope estimates to compute the average slope for the whole sample and its standard error. The third row in Table II shows that the Fama–MacBeth method yields a coefficient of -1.4 and a standard error of 0.34. (Weighting the monthly slope coefficients by the number of observations in each month yields a coefficient of -0.76 with a standard error of 0.23.) Therefore, when we take into account the possibility of correlated residuals in our estimation of standard errors, we also reject, at any conventional significance level, the hypothesis that the slope of the fee-performance relation is one. Moreover, our estimate of the slope coefficient is negative and significantly different from zero at the 5% (1%) significance level when we use clustered (Fama–MacBeth) standard errors to perform our tests.

A potential problem with our results is that estimated alphas contain funds' true abnormal performance, but they also contain estimation error from two sources, the residuals of the performance attribution model (1) and estimation error in the realized risk premium. Estimation error in the dependent variable in (2) may affect inference in several ways. First, this estimation error increases the variance of the residuals and thus the standard errors of parameter estimates. Therefore, estimation error in alphas decreases the likelihood of finding a significant relationship between alphas and expense ratios. Second, if extreme alphas are more likely to include a large estimation error and have a large influence on the estimated coefficient, then estimation error in alphas may affect the estimated slope coefficient. More generally, even in the absence of estimation error in alphas, our results could be driven by a relatively small number of funds with extreme alphas or expenses.

To explore the monotonicity and linearity of the fee-performance relation, Figure 1 shows conditional expected risk-adjusted performance as a nonparametric function of expenses. We estimate the conditional expectation by using the Nadaraya–Watson estimator with a Gaussian kernel (see, for example, Härdle (1990)). We also plot conditional expected alphas as a linear function of expenses as implied by the estimated OLS coefficient of the linear regression. To account for time effects, we de-mean expenses and risk-adjusted performance by subtracting the month's average. For values of expenses below the sample's 99th percentile, Figure 1 shows that before-expense risk-adjusted performance decreases monotonically with expenses and that the relation may be well described by a linear function. However, for funds with expenses in the top sample percentile, the relation appears far from monotonic, although the large confidence intervals in this low density region suggest that inference on the mean risk-adjusted performance of very expensive funds is problematic. Therefore, Figure 1 shows that the presence of some funds with both extreme expense ratios and extreme risk-adjusted performance does not seem to affect the OLS slope coefficient.

Finally, although all funds are affected by estimation error in betas, the standard errors of estimated betas may vary across funds in systematic ways. Thus, although White (1980) and clustered standard errors already account for heteroskedasticity, we also run several generalized least squares regressions, obtaining results (see the Internet Appendix) that are essentially identical to those obtained by OLS.

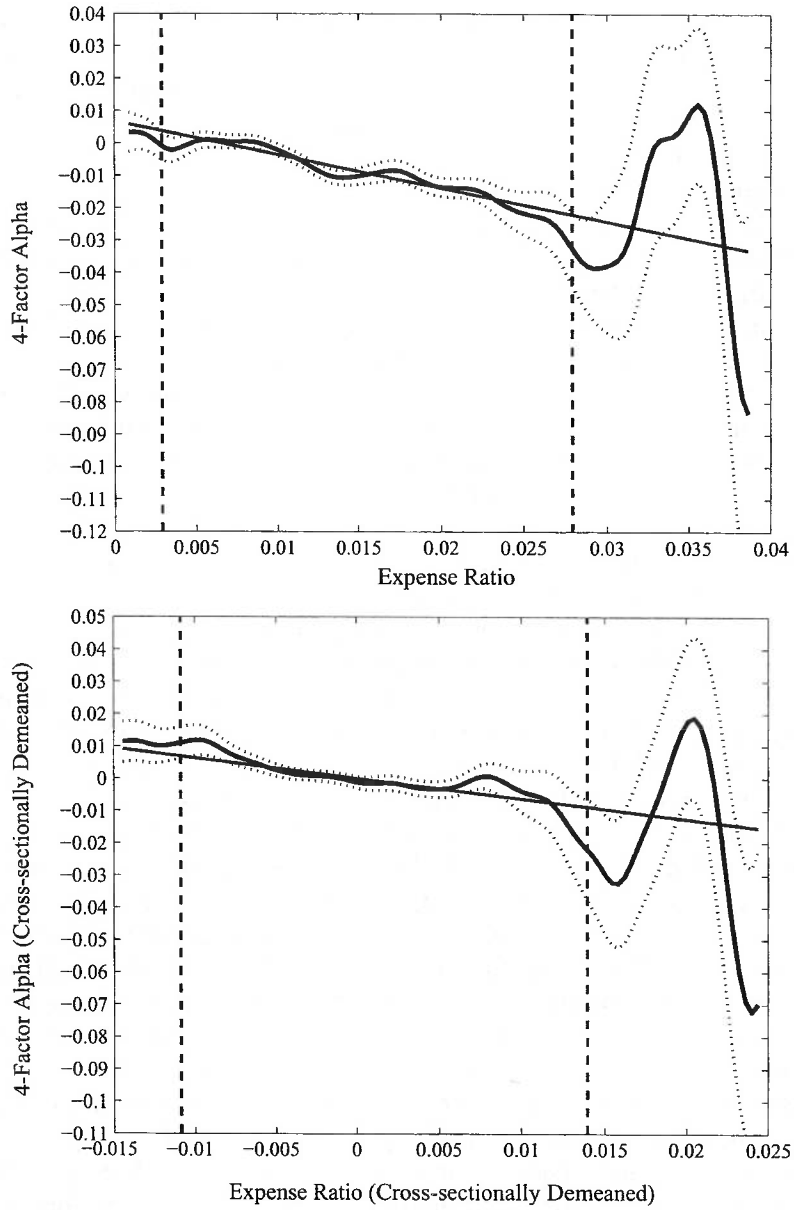


Figure 1. Nonparametric Regressions. The figure shows mutual fund conditional expected risk-adjusted performance as a nonparametric function of expense ratios (solid thick line). Risk-adjusted performance in month t is estimated as the difference between the fund's monthly return in month t and the product of betas and the factor realizations for that month using Carhart's four-factor model. Both monthly risk-adjusted performance and expense ratios are annualized. The conditional expectation has been estimated using the Nadaraya–Watson estimator with a Gaussian kernel. Dotted lines show upper and lower bounds of the 95% pointwise confidence interval. The figure also plots conditional expected alphas as a linear function of expenses as implied by the estimated OLS coefficient of the linear regression (solid thin line). Dashed vertical lines correspond to the 1st and 99th percentiles of the expense ratio sample distribution. The bottom panel displays the results obtained when we de-mean expenses and risk-adjusted performance by subtracting the month's sample average.

To further account for potential estimation error in alphas stemming from the estimation of betas, we also report results obtained using other performance evaluation models. The last row of Table II shows the results that we obtain when we estimate alpha using the Fama–French three-factor model. The estimated coefficient is -0.2 with a standard error of 0.26. Therefore, we can reject the null hypothesis that the slope of the fee-performance relation is one at the 1% significance level. However, the estimated coefficient is not significantly different from zero, in contrast to the results obtained with Carhart's four-factor alpha. This finding suggests that on average, more expensive funds exhibit greater exposure to the momentum factor. Further, in addition to the three- and four-factor unconditional models, we use several conditional versions of Carhart's four-factor model (see, for example, Ferson and Schadt (1996) and Kosowski et al. (2006)) and obtain results consistent with those from the unconditional model. In the Internet Appendix, we describe in detail the specification and results of the conditional models.

Another possible concern about our results is that they may be due to the influence of funds with small market share. Such funds may exhibit both low performance and high expense ratios. However, our requirement that funds have at least 48 months of return information to be included in the sample already filters out the effect of unsuccessful funds that are terminated before reaching that threshold. To evaluate the influence of small funds that have survived for at least 5 years, we reestimate equation (2) for a sample that excludes observations with relatively low values of assets under management in each month. Panel A of Table III shows that the negative relation between expense ratios and before-expense risk-adjusted performance holds when the lowest size decile is excluded each month. Although the estimated coefficient is significantly different from one at the 1% significance level, it is smaller in absolute value than our estimate for the whole sample and only marginally significantly different from zero. Excluding further deciles leads to similar coefficients and standard errors (with coefficients that are either not statistically significant or only marginally so). Thus, fund size appears to play a role in explaining the relation between risk-adjusted performance and fund expenses.

In the analysis above, we consider expense ratios as the only explicit cost of delegated portfolio management. However, investors often pay loads at the time of purchasing and/or redeeming mutual fund shares. Hence, the previous regressions could be capturing a negative relation between performance and a specific component of total fund share ownership cost, but not necessarily a negative relation between performance and the total fees paid by investors. In particular, if more expensive funds (when only expenses are considered) charged lower loads, then after-fee performance (when all fees are considered) could still be equalized across funds.

One way to circumvent this problem is to focus exclusively on funds for which annual operating expenses account for 100% of all fees. In Table III (Panel B), we estimate equation (2) for no-load funds only. The estimated slope coefficient of -0.91, which is significant at the 1% level, indicates that total ownership cost and performance are negatively correlated for no-load funds.

Table III

Regressions by Subsamples

The table shows estimated slope coefficients for the OLS regression of funds' monthly before-fee risk-adjusted performance on monthly fees in the period from January 1962 to December 2005 for Panels A to C, and the period from January 1992 to December 2005 for Panel D. Betas are estimated using Carhart's four-factor model with a 5-year estimation period. Risk-adjusted performance in month t is the difference between the fund's monthly before-expense return and the product of betas and the factor realizations in t. Monthly fees are defined as the annual expense ratio divided by 12, except for Panel B, where monthly fees are annual expense ratios divided by 12 plus the sum of front-end and back-end loads divided by the assumed holding period in months. In Panel A the sample does not include for each month the decile of fund-month observations with the lowest total net assets among all actively managed retail funds. In Panel B No-Load Funds are defined as those charging no front- or back-end loads. All regressions include dummies for months. Standard errors (in parentheses) are clustered by time. Adjusted R^2 statistics are reported in percentage. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively. Superscripts a, b, and c denote that the null hypothesis of a unit coefficient is rejected at the 10%, 5%, and 1% significance levels, respectively.

Subsample	Coefficient	Adj. R^2	Obs.
Panel A: Effect of Small Funds			
Deciles 2–10	$-0.4602^{*,c}$	10.28	225,450
	(0.2693)		
Panel B: Other Fees			
No-load funds	$-0.9099^{***,c}$	8.51	81,323
	(0.3220)		
Load funds (2-year holding period)	$-0.2660^{***,c}$	11.22	150,148
	(0.1003)		
Load funds (7-year holding period)	$-0.5186^{*,c}$	11.22	150,148
	(0.2823)		
Panel C: Regressions by Subperiods			
1967–1976	-0.9938^{b}	14.26	16,504
	(0.9190)		
1977–1986	$-0.9581^{*,c}$	8.81	26,591
	(0.5625)		
1987–1996	$-0.8129^{***,c}$	5.65	39,567
	(0.2560)		
1997–2005	-0.5384^{c}	10.43	149,724
	(0.3347)		
Panel D: Regressions by Investment Objective			
Aggressive growth funds	-0.4304^{a}	17.44	13,419
	(0.7244)		
Growth MidCap funds	-0.0183^{a}	27.99	15,366
	(0.6181)		
Growth and income funds	$-0.5840^{***,c}$	12.79	39,221
	(0.1619)		
Growth funds	$-0.6557^{**,c}$	8.48	70,277
	(0.2757)		
Small company growth funds	$-0.6931^{*,c}$	21.88	35,065
	(0.3974)		

Because load funds constitute two-thirds of the sample, we also estimate the relation between performance and a measure of total fund ownership cost for these funds. Following Sirri and Tufano (1998), we compute total annual ownership costs by adding annuitized total loads (total loads divided by the number of years, τ, that investors keep their money in the fund) to annual expense ratios. Although previous studies typically set $\tau = 7$, redemption rates for equity funds for more recent periods suggest a shorter average holding period in the range of 2.5 to 5 years. Therefore, we perform the analysis for $\tau = 2$ and 7 years. Because the analysis in Table III is conducted at a monthly frequency, the independent variable is total monthly ownership cost, defined as total annual ownership cost divided by 12. Panel B of Table III shows that total ownership cost is negatively and significantly associated with before-fee risk-adjusted performance for both holding periods. Further, the unit slope hypothesis is rejected at any conventional significance level.

As a final test of the robustness of our regression results, we estimate equation (2) for different subperiods and mutual fund categories. Panel C of Table III shows that the perfectly competitive equilibrium condition is clearly violated in all subperiods considered. Moreover, the relation between before-fee risk-adjusted performance and expenses is negative in all the subperiods, although not significantly different from zero in the 1967 to 1976 and 1997 to 2005 subperiods. Lack of significance in the 1967 to 1976 subperiod could be due to the relatively low number of observations, which results in a large standard error for the slope coefficient. However, failure to reject the null hypothesis of a zero coefficient in the last subperiod appears to happen because the fee-performance relation becomes flatter in the last years of our sample: Although the estimated slope coefficient lies between -0.81 and -0.99 in the pre-1997 years, it is -0.54 in the last subperiod considered.

Finally, for the 1992 to 2005 period, for which the classification is detailed and consistent, we divide the sample into subsamples according to the Standard & Poor's detailed objective code as reported by CRSP, and then run the regression for each subsample. Panel D of Table III shows that expense ratios are negatively related to performance for all five investment objectives, although the relation is not statistically significant for Aggressive Growth and Growth MidCap funds. For these investment objectives, the unit slope hypothesis can be rejected, but only at the 10% significance level. When we replace expense ratios with total ownership cost with $\tau = 2$ or 7 years, we obtain results that are similar to those of Panel D (see the Internet Appendix for results).

IV. Explaining the Relation between Fees and Performance

The negative relation between before-fee performance and fees that we uncover in the previous section is at odds with the intuitive expectation that fees should, at least to some extent, reflect the value that funds create for investors. In this section, we set forth and test different explanations for this apparently anomalous relation.

A. Cost-Based Explanations

According to the first explanation, fees simply reflect the costs of operating the fund. If low costs are associated with better before-fee risk-adjusted performance, then a univariate regression would result in a negative relation between fees and performance.

Fund performance could be positively associated with fund costs if higher costs reflect higher salaries to attract more talented managers or a larger investment in research tools, but there are also arguments for a negative correlation between costs and performance. For instance, there might be economies of scale that lower operating costs for larger funds. In addition, larger size may be associated with better performance if a fund's size reflects its past performance, and performance is persistent. Similarly, older funds might benefit from learning economies, which could be passed on to investors in the form of lower fees. If fund longevity is related to good performance, as would be the case if low-performance funds were more likely to close down, we could observe a negative relation between costs and performance. Finally, higher managerial skill may be associated with both better investment decisions and more efficient management of fund operations, which would translate into lower operating costs.

B. Strategic Explanations

The second explanation views the negative relation between before-fee performance and fees as the result of strategic fee-setting by mutual fund management companies or other service providers to the fund. One such explanation has been proposed and empirically tested for money market mutual funds by Christoffersen and Musto (2002). On the basis of empirical studies on mutual fund flows (e.g., Sirri and Tufano (1998)) and survey data on mutual fund investors' behavior (Capon, Fitzimmons, and Prince (1996), Alexander, Jones, and Nigro (1997)), Christoffersen and Musto argue that mutual fund investors differ in their performance sensitivity. They also argue that funds with a worse performance history will have a less performance-sensitive clientele because the performance-sensitive investors will have fled those funds following bad performance. Therefore, funds with a greater proportion of performance-insensitive investors will charge higher fees because for these funds the reduction in after-fee performance caused by an increase in fees will not translate into a large flow of money out of the fund. It follows that funds with bad past performance will find it optimal to charge higher fees. Christoffersen and Musto's explanation can be tested by using a measure of the performance sensitivity of each fund's flows.

Gil-Bazo and Ruiz-Verdú (2008) provide a related strategic explanation for the negative relation between before-fee performance and fees. These authors develop an asymmetric information model of the mutual fund market in which mutual funds differ in their expected performance and investors differ in their performance sensitivity. Gil-Bazo and Ruiz-Verdú show that competition for the money of performance-sensitive investors leads to an equilibrium in which funds that expect to earn higher returns ("good" funds) reduce their fees up to the point at which they effectively price funds that expect lower returns ("bad" funds) out of the performance-sensitive segment of the market. Good funds are able to price bad funds out of the market because the revenues of management companies are determined as a fraction of assets under management. Therefore, for any given fee (expressed as a fraction of asset value), good funds, which can be expected to achieve a larger increase in the value of their assets, will earn higher expected revenues. As a result, there is a fee level at which good funds break even in expectation, and low-performance funds incur an expected loss. Unable to compete for performance-sensitive investors, bad funds raise their fees to extract rents from performance-insensitive investors. Gil-Bazo and Ruiz-Verdú's predictions can also be tested by using a measure of a fund's risk-adjusted expected performance.

A related explanation for our results is that low-performance funds incur higher marketing costs and those costs are passed on to investors in the form of higher fees. If low-performance funds target performance-insensitive investors and these investors purchase mutual fund shares mostly through brokers, then low-performance funds will incur higher marketing costs than funds sold through more direct distribution channels (e.g., from mutual fund supermarkets or directly from the management company). Bergstresser, Chalmers, and Tufano (2009) provide evidence that supports this hypothesis. These authors report that on average funds sold through the direct or fund supermarket channels perform better and have lower fees than those sold through the broker channel. The marketing costs of low-performance funds could also be higher if intermediaries have to be compensated for the higher effort or potential loss of reputation associated with selling low-performance funds.[9] Alternatively, if unsophisticated investors are more responsive to advertising and low-performance funds target those investors, then low-performance funds will spend more on advertising because the marginal return of their advertising investment will be higher. These strategic marketing explanations imply that underperforming funds will have higher marketing fees, an implication that we can test with our data.

Both Gil-Bazo and Ruiz-Verdú's (2008) explanation and the strategic marketing hypothesis assume that fund management companies form expectations about the future performance of the funds they manage, and that they condition their funds' fees on those expectations. This assumption seems reasonable because management companies are able to observe all publicly available information about funds' portfolio choices and returns, they have access to a wealth of data not available to outsiders (such as high-frequency data on portfolio holdings), and they have the skills to analyze all that information. Further, management companies themselves may strongly influence the performance of their funds because they decide how to allocate scarce resources (staff, research analysis, underpriced IPOs) among them (Gaspar, Massa, and Matos

[9] We thank an anonymous referee for suggesting this possibility.

(2006), Guedj and Papastaikoudi (2005)). However, a limitation of these hypotheses is that they do not explain why, rather than adjusting fees, management companies do not try to improve the expected performance of their funds by replacing underperforming managers or changing their investment strategies. In particular, funds with low expected performance could be turned into closet indexers, guaranteeing a level of performance close to the benchmark. Addressing these limitations is beyond the scope of this paper, but allowing for managers' replacement or changes in investment strategies may not substantially change the predictions of the strategic explanations. First, not all companies will be able to hire only the managers with the highest expected performance, and mutual funds with less-than-top managers may be able to survive, at least in the medium run, especially in the presence of unsophisticated investors. Second, even if closet indexing guarantees that returns do not fall too much below a fund's benchmark, closet indexers will still underperform funds that have managers who are able to generate positive alphas. Further, a strategy of closet indexing also entails some less-obvious costs. In particular, sophisticated investors will leave a fund that they identify as a closet indexer because closet indexers are dominated in terms of after-fee returns by index funds.

Another potential limitation of the strategic explanations is that although many studies document the existence of a significant pool of unsophisticated investors, it is an open question whether unsophistication can persist in the medium or long run. In particular, cheaper or better-performing funds may want to educate performance-insensitive investors to avoid expensive funds. However, Gabaix and Laibson (2006) show that firms may not have incentives to educate investors to avoid "shrouded" costly attributes (product attributes that are hidden by firms, even though they could be almost costlessly revealed). Even though most of Gabaix and Laibson's analysis focuses on add-ons (product attributes that the consumer can substitute away at a cost), the authors also discuss factors that might limit firms' incentives to educate unsophisticated investors about unavoidable shrouded costs.

C. Fund Governance

The strategic explanations discussed earlier implicitly assume that mutual fund fees are set by management companies so as to maximize fee revenues. However, U.S. mutual funds are legal entities that are independent of the companies managing their portfolio. Control over the fund is delegated by fund shareholders to a board of directors (or trustees), which is responsible for contracting the management of the fund's portfolio with a management company. Thus, the management fee is not set unilaterally by the management company; rather, it is negotiated with the board of directors. Similarly, the fund's directors negotiate the fees paid to other service providers, such as distributors or transfer agents. Boards of directors have the fiduciary duty to ensure that those fees reflect the value for fund investors of the services they are paying for.

Despite legal provisions imposing rigorous governance requirements on mutual funds, there remain important conflicts of interest that may interfere with fund directors' fiduciary duty. For instance, management companies select the members of a fund's initial board of directors. In practice, contract renegotiations and changes in the fund's management company are infrequent (Kuhnen (2005), Warner and Wu (2006)), suggesting that directors' interests may be more aligned with those of management companies than with those of fund investors.

Prior research is not fully conclusive as to whether certain mutual fund governance structures are able to mitigate these conflicts of interest and lead to lower fees. Although Tufano and Sevick (1997) provide evidence for 1992 that funds with smaller boards and funds with boards that have a higher fraction of independent directors have lower fees, more recent studies (Meschke (2007), Ferris and Yan (2007)) obtain mixed results. We hypothesize that the boards of better-governed funds will approve "fair" or "reasonable" fees in fulfillment of their fiduciary duty. This does not necessarily mean lower fees (although better governance could also result in lower fees), but rather that fees are more in line with the fund's performance. Therefore, the relation between fees and performance should be positive, or at least flatter, for better-governed funds. Similarly, boards of higher quality may resist more strongly any attempts by management companies and other service providers to charge higher fees in funds with more performance-inelastic investors. If this were the case, the relation between fund fees and performance sensitivity would be flatter for better-governed funds. Although with limitations imposed by the nature of our data, we test these hypotheses using Morningstar's board quality grade as a measure of fund governance.

D. Empirical Strategy

To test the empirical validity of the proposed explanations for the negative relation between fees and before-fee risk-adjusted performance, we investigate how fees vary with fund characteristics, flow-to-performance sensitivity, and performance. We assume that fund i's fee at time t, f_{it}, is a linear function of a vector $\mathbf{x}_{it-1}$ of lagged values of variables that are likely to determine the fund's operating costs, the performance-sensitivity of the fund's flows, S_{it}, and the fund's expected before-fee performance in period t, α_{it}:

$$f_{it} = \gamma' \mathbf{x}_{it-1} + \lambda_S S_{it} + \lambda_\alpha \alpha_{it} + \nu_{it}, \tag{3}$$

where ν_{it} is a generic error term. Because data on most variables are available yearly during most of the sample period, the time index t in equation (3) refers to calendar years.

To test the potential effects of fund governance on fund fees, we also estimate an extended version of equation (3) that allows both the intercept and the coefficients on performance and performance sensitivity to depend on board quality.

We build on the literature on mutual fund fee determinants, which mostly considers fund fees as a reflection of operating costs, to select the variables that may influence the costs of operating a fund.[10] For every fund-year observation we consider the following variables: *size*, which we define as the log of the year-end total net asset value; *age*, computed as the log of the number of years since the fund's organization; *size of the complex* and *number of funds in the complex*, which we define as the log of the total net asset value of the funds managed by the company that manages the fund, and the total number of funds managed by that company, respectively; reported annual *turnover*; *volatility*, computed as the standard deviation of the fund's monthly returns in the year; and dummy variables for the fund's *investment objective*. We also include a dummy variable to identify single-class load funds and dummies for the main share classes. Doing so makes it possible for us to correct for the potential distortions induced by using a homogeneous holding period for all funds because we can expect investors with different holding periods to select different share classes. We include time dummies in all regressions.

We use $\widehat{\alpha}_{it}$, which we define as the sum of estimated monthly alphas in year t, as our proxy for the fund's expected before-fee performance.[11] Estimated alpha ($\widehat{\alpha}_{it}$) is a good measure of expected performance (α_{it}) as long as the measurement error in $\widehat{\alpha}_{it}$ is not correlated with the level of fees. If no such correlation exists, the result of including estimated, rather than expected, performance as a regressor reduces to the well-known attenuation bias in the presence of measurement error. Thus, the performance coefficient estimates are likely to be biased toward zero.

To obtain a measure of the flow-to-performance sensitivity, S_{it}, we proceed in two steps. First, we estimate a model of money flows into mutual funds. Based on prior studies of fund flow determinants, we allow the sensitivity of flows to past performance to depend on both the level of performance and fund characteristics. In the second step, we estimate flow-to-performance sensitivity for each fund and year as the first derivative of conditional expected flow with respect to the previous year's performance. The Appendix provides a detailed description of the construction of this variable.

E. The Determinants of Fund Fees

To investigate the relation between before-fee performance and the total fees paid by investors, we first use total ownership cost as our dependent variable. We compute total ownership cost as the expense ratio plus total loads divided by seven. Since we consider a 7-year holding period, to account for usual

[10] Different aspects of mutual fund fee determination have been studied, among others, by Ferris and Chance (1987), Tufano and Sevick (1997), Latzko (1999), Malhotra and McLeod (1997), Chalmers, Edelen, and Kadlec (2001), Luo (2002), Deli (2002), and Golec (2003).

[11] In the remainder of the paper, we focus on unconditional Carhart's alpha exclusively. We use $\widehat{\alpha}_{it}$ as a measure of the alpha expected by the manager of fund i at the beginning of period t under the assumption that fees are set at the beginning of period t. If fees were set in the middle of period t, our measure of expected performance would thus aggregate performance observed prior to setting fees with expected performance. We have estimated the fee equation using $\widehat{\alpha}_{it+1}$ as a measure of expected returns and obtained identical results.

Table IV

Mutual Fund Fee Determinants

The table reports estimated coefficients for yearly regressions of funds' fees on selected fund characteristics in the 1993 to 2005 period. The dependent variable in columns (1) to (6) is total annual ownership cost (*TOC*), computed as total loads divided by seven plus the annual expense ratio. In columns (7) and (8), the dependent variable is marketing fees (*Mark.*), defined as total loads divided by seven plus 12b-1 fees, and nonmarketing fees (*N-Mark.*), computed as the expense ratio minus 12b-1 fees, respectively. Back-end loads are assumed to be zero for share classes B and C. The coefficients in columns (1) to (3) and (7) to (8) are estimated by pooled OLS. Column (4) reports estimated coefficients for a regression with management company fixed effects. In column (5) all variables are asset-weighted averages at the management company level. Column (6) reports estimated coefficients for a regression with fund fixed effects. The size of the management company and the number of funds in the management company are denoted by *Co. Size* and *# funds*, respectively. σ_t is the standard deviation of the fund's monthly returns in year *t*. S_t denotes the slope of the estimated flow-to-performance relation. $\widehat{\alpha}_t$ is the year *t* four-factor alpha. All regressions include year dummies and dummy variables for the different investment objectives and share classes. All fees are expressed in bp. The table also reports robust standard errors (in parentheses), which are clustered by fund in columns (1) to (3) and (6) to (8), and by management company in columns (4) to (5). The total number of observations and the adjusted R^2 of the regression (in percentage) are reported at the bottom of the table. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	TOC						Mark.	N-Mark.
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Size_{t-1}	−9.38***	−9.77***	−9.76***	−7.04***		−7.25***	−6.35***	−3.40***
	(0.70)	(0.82)	(0.82)	(1.02)		(0.80)	(0.60)	(0.54)
Age_{t-1}	−4.96***	−8.65***	−8.63***	−12.05***	−0.18	−9.91***	−2.99***	−5.61***
	(1.55)	(1.73)	(1.73)	(1.69)	(0.19)	(0.00)	(1.10)	(1.29)
Co. Size_{t-1}	−1.06	−1.17	−1.14	−3.75**	−12.82***	−1.39*	7.46***	−8.64***
	(0.73)	(0.81)	(0.80)	(1.75)	(1.58)	(0.84)	(0.64)	(0.55)
$\text{\#funds}_{t-1}$	−0.09	−0.14	−0.14	0.16	1.44***	−0.05	−0.66***	0.52***
	(0.09)	(0.10)	(0.10)	(0.19)	(0.53)	(0.10)	(0.09)	(0.06)
Turnover_{t-1}	3.29***	3.27***	3.19***	0.14	9.84***	1.93***	−0.70	3.92***
	(0.73)	(0.86)	(0.86)	(1.04)	(2.34)	(0.82)	(0.70)	(0.60)
σ_{t-1}	42.53***	34.70***	28.18**	13.09*	58.53**	−12.22***	17.89**	16.33**
	(10.45)	(11.49)	(11.34)	(7.62)	(27.78)	(6.01)	(8.73)	(7.92)
$\widehat{\alpha}_t$	−14.65***	−18.56***	−19.48***	−11.74**	−39.16**	−5.10*	−11.79***	−6.94*
	(5.33)	(5.64)	(5.68)	(5.00)	(17.77)	(2.92)	(4.41)	(3.67)
$\widehat{\alpha}_t^2$			75.84***					
			(27.38)					
S_t		−7.06***	−7.01***	−3.81***	−6.54***	−0.70	−1.18*	−5.87***
		(1.11)	(1.11)	(0.96)	(2.10)	(0.55)	(0.70)	(0.82)
Obs.	12,709	10,290	10,290	10,290	1,580	10,290	10,284	10,353
Adj. R^2	52.11	55.19	55.23	74.67	54.83	94.22	59.17	42.69

practices we assume that effective back-end loads are zero for classes B and C shares.

Column (1) in Table IV presents the results of estimating equation (3) without the performance-sensitivity measure. If the negative relation between expected performance and fees were the consequence of the omission of variables (for example, size, age, or turnover) that are likely to determine operating costs and are related to performance, then we would expect the coefficient on expected performance to change sign, or at least to become statistically insignificant once we include these variables in the regression. As column (1) shows, this is not the case: The coefficient on expected performance remains negative and significantly different from zero. Therefore, cost-based arguments cannot explain why fees and performance are negatively related.

Column (2) reports the results of estimating the full model (3), in which we include both performance and performance-sensitivity as regressors. The negative (and statistically significant at the 1% level) coefficient on performance sensitivity suggests that equity mutual funds strategically exploit a low elasticity of demand with respect to net performance to increase their fees. Therefore, our results extend the findings of Christoffersen and Musto (2002), which were obtained for a cross-section of money market mutual funds, to the market for actively managed equity mutual funds, for a much larger sample, and with a more precise measure of performance sensitivity. However, the inclusion of performance sensitivity does not eliminate the negative association between expected before-fee risk-adjusted performance and fees. On the contrary, the estimated coefficient in column (2) is not only negative and statistically significant, but higher (in absolute value) than the estimated coefficient in column (1). Thus, elasticity of demand appears to be an important determinant of fees, but does not in itself explain why underperforming funds set higher fees.

To clarify the economic significance of these results, the estimated coefficients in column (2) imply that a one-standard deviation increase in annual before-fee risk-adjusted performance is associated with a decrease of 1.38 basis points in annual total ownership cost, while a one-standard deviation increase in performance sensitivity is associated with a 4.22-basis point decrease in fees per year. To put these figures in perspective, increases of one-standard deviation in fund volatility and turnover are associated with increases in total ownership cost of 1.99 and 4.79 basis points, respectively, while an increase of one standard deviation in fund size is associated with a reduction in total ownership cost of 23.21 basis points.

To account for possible nonlinearities, in column (3) we include alpha squared in the regression. The associated positive coefficient suggests that the relation between alpha and fees becomes flatter for funds with higher alphas, although it remains negative for plausible values of alpha.

These results indicate that worse-performing funds and those whose investors have a lower performance sensitivity charge higher fees. Because management companies typically manage many funds, the results could be due to differences between management companies, differences within management companies, or a combination of both. Column (4) of Table IV reports the results when we estimate equation (3) including a management company fixed effect to capture time-invariant differences between management companies. The estimated coefficients indicate that differences in fees within management companies are negatively related to differences in alpha or performance sensitivity. Thus, the results are consistent with management companies strategically setting the fees of their different funds to match the funds' expected performance or the performance sensitivity of their investors.[12] The large increase in the adjusted R^2 of the regression with respect to the R^2 obtained in columns (1) to (3) also suggests that there are management company characteristics beyond the company's total size and number of funds that are related to fees. (The joint hypothesis that all management company fixed effects are zero can be rejected at any reasonable significance level.) We also generate observations at the level of the management company by taking asset-weighted averages of all variables, except for complex size and the number of funds in the complex, for which we retain the management company totals. Column (5) of Table IV reports the results of estimating equation (3) at the management company level. For each management company we compute the asset-weighted average of each variable and exclude management companies for which total net assets of funds with information on the variable are less than 75% of the total net assets of the management company. We also exclude management companies with more than one-third of assets under management in index or institutional funds or more than 10% of assets in funds categorized as outliers. The estimated coefficients for alpha and sensitivity indicate that the results obtained at the fund level extend to the management company level.

To evaluate whether marketing or nonmarketing fees are responsible for the negative relation between total fees and performance, we reestimate equation (3), replacing the dependent variable with marketing and nonmarketing fees, alternatively. We define marketing fees as the sum of front- and back-end loads divided by seven (except for B and C share classes, for which we only add front-end loads) and 12b-1 fees. We define nonmarketing fees as the expense ratio minus 12b-1 fees. Columns (7) and (8) of Table IV show that both marketing and nonmarketing fees are negatively related to flow-performance sensitivity and to risk-adjusted performance.[13] Therefore, the results are consistent with the three strategic explanations discussed in Section IV.B. In particular, the results obtained for marketing fees suggest that marketing variables, such as distribution channel or advertising expenditures, play a significant role in determining mutual fund fees and contribute to explaining the negative fee-performance relation.

Although our results are consistent with the strategic fee-setting hypotheses, the negative relation between before-fee performance and fees could also be caused by unobserved fund characteristics that have correlations of opposite signs with fees and before-fee risk-adjusted performance. To control for time-invariant unobserved heterogeneity and to shed some light on the determinants of fee changes, we also estimate the fee equation (3) with fund fixed

[12] Our estimates with management company fixed effects might also pick up longitudinal variation at the management company level. To isolate the cross-sectional variation, we also estimate yearly cross-sectional regressions with management company fixed effects and obtain coefficient estimates and standard errors (see the Internet Appendix) for the whole sample period, using the Fama–MacBeth procedure. The resulting coefficients for alpha and performance sensitivity are negative, statistically significant, and only slightly smaller (in absolute value) than the corresponding coefficients obtained without management company fixed effects.

[13] Given the large incidence of funds with zero marketing fees, we also estimate a Tobit model and obtain very similar coefficients and standard errors (results available in the Internet Appendix).

effects. Column (6) of Table IV shows that the signs of the coefficients on performance and performance sensitivity remain negative in the fixed effects specification. Thus, funds appear to alter their fees over time in response to changes in performance in the same direction found in the absence of fund fixed effects.[14] However, the coefficients on performance and performance sensitivity are smaller in absolute value than those in column (2) and the latter coefficient is not statistically different from zero at conventional significance levels. The smaller value and reduced statistical significance of the estimated coefficients on performance and performance sensitivity in the fund fixed effects specification could be due to fees being largely a matter of long-term strategy, or to the existence of adjustment costs that create substantial inertia in fees. The differences between the specifications could also be due to the fact that our performance and performance sensitivity measures are estimated variables that are likely to contain substantial measurement error. This measurement error may have a larger impact on coefficient estimates when we use only time-series variation to estimate them. For example, in the extreme case in which fund managers' skills are constant over time, the entire time-series variation in alphas would be due to estimation error. At the same time, the cross-sectional variation in estimated performance would allow us to pick up at least part of the effect of the true alpha.

We note that there is a possible alternative explanation for the negative relation between fees and before-fee performance that we do not explicitly consider in our analysis. Higher fees could be paying for better tax management or other fund services, such as check writing, web or telephone services, or better shareholder statements, that compensate investors for differences in after-fee performance. Although our data do not enable us to rule out this explanation, there are reasons to believe that its ability to explain our findings is limited. First, the explanation is valid only if the value of these services is negatively related to fund performance. Moreover, if better fund services or more efficient tax management fully compensated investors for lower after-fee performance, then there would be no reason to expect investors' money to flow into funds with higher after-fee performance. The explanation also finds little support from the empirical evidence on both the perceived and actual value of fund services, which seem to be limited (Capon, Fitzimmons, and Prince (1996), Elton, Gruber, and Busse (2004), Christoffersen, Evans, and Musto (2005), Bergstresser, Chalmers, and Tufano (2009)). The fact that we find a negative relation between fees and performance when we include fund fixed effects also casts doubt on the plausibility of this interpretation, unless the level of services changes over time in a direction opposite to performance.

F. Fund Governance and Fees

To analyze the role played by fund governance in the determination of fees, we create dummy variables corresponding to each one of the Morningstar board

[14] In tests available in the Internet Appendix, we also estimate the fee equation in differences, as well as logit and probit regressions for the probability of fee changes, and obtain similar results.

quality grades. Because only one fund in our sample received Morningstar's lowest board quality rating, we create a single dummy for the Very Poor and Poor categories, which we refer to as "Poor." We also create interaction variables of board-quality dummies with before-fee risk-adjusted performance and with our flow-to-performance sensitivity measure.

In Table V, we present our regression results. Because of the limitations of our governance subsample, discussed in Section I, these results should be interpreted with caution. First, the fact that our governance information corresponds to the end of the sample period makes a causal interpretation of the governance coefficients problematic and implies that the results may be subject to survivorship bias. Further, the governance subsample is not a random sample of the population. As we describe in Section I, funds in the governance subsample tend to have lower-than-average fees and higher-than-average performance. Keeping these caveats in mind, the results in columns (1) to (5) of Table V provide some support for the hypothesis that better board quality leads to fees that better reflect the value generated by mutual funds, and that better board quality also limits the ability of management companies to extract rents from performance-insensitive investors.

Column (1) presents the results of estimating the fee equation when we exclude the performance and performance sensitivity terms, but allow the intercept to depend on the board quality grade. In the table, we omit the dummy variable for the Poor grade, so coefficients associated with Fair, Good, or Excellent grades represent differences with respect to funds with a Poor grade. Funds with Good and Excellent grades are associated with lower total ownership costs than Poor and Fair funds, although the difference between the coefficients associated with Excellent and Poor grades is not statistically different from zero. Column (2), which reports the results of estimating the full model, also implies that for most values of before-fee alpha or performance sensitivity in the sample, funds with Good or Excellent grades are predicted to be cheaper, all other things equal, than funds with Poor or Fair grades (although this conclusion cannot be directly inferred from the table because the difference between the fees of two funds identical in every respect except for their board quality grade is not just captured by the intercept, but depends on the levels of alpha and performance sensitivity). Thus, columns (1) and (2) suggest that better board quality is associated with lower total ownership costs. However, although in no case are the total ownership costs of a fund with a higher board quality grade significantly higher than those of a fund with a lower grade, some of the estimated differences are not statistically significant and the levels of the estimated coefficients are not strictly monotonic in board quality.

If better board quality brings fees more in line with performance, then the coefficient on before-fee performance should be negative for poorly governed funds and the interaction coefficients should be positive and increasing in board quality. The estimated coefficients on performance in column (2) are consistent with this hypothesis. The coefficients on performance sensitivity point in the same direction, but in this case none of the differences is statistically significant at conventional significance levels.

Table V

Fund Governance and Fees

The table reports estimated coefficients for the pooled OLS regression of funds' fees on selected fund characteristics in the 1993 to 2005 period, except for columns (6) and (7), where the sample period is 2003 to 2005. In columns (1) and (2) the dependent variable is fund total annual ownership cost (*TOC*), computed as total loads divided by seven plus the annual expense ratio. In column (3) the dependent variable is the asset-weighted average at the management company level of total annual ownership costs. All regressors in column (3) are also asset-weighted averages at the management company level of the corresponding fund-level variables, except for the size of the management company (*Co. Size*) and the number of funds in the management company (*# funds*), which are management company totals. In column (4) the dependent variable is marketing fees (*Mark.*), defined as total loads divided by seven plus 12b-1 fees. Back-end loads are assumed to be zero for share classes B and C. In columns (5) and (7) the dependent variable is nonmarketing fees (*N-Mark.*), computed as the expense ratio minus 12b-1 fees. Column (6) shows results for management fees (*Mgmt.*). σ_t is the standard deviation of the fund's monthly returns in year t. $\widehat{\alpha}_t$ is year t's four-factor alpha. S_t denotes the slope of the estimated flow-to-performance relation. *Fair*, *Good*, and *Excell.* are dummy variables that take a value of one if the observation has Fair, Good, or Excellent board quality grade, respectively, and zero otherwise. All regressions include year dummies and dummy variables for the different investment objectives and share classes. All fees are expressed in bp. The table also reports robust standard errors (in parentheses) clustered by fund in all columns except column (3), in which they are clustered by management company. The total number of observations and the adjusted R^2 of the regression (in percentage) are reported at the bottom of each column. *, **, *** indicate statistical significance at the 10%, 5%, and 1% levels, respectively.

	TOC (1)	TOC (2)	TOC (3)	Mark. (4)	N-Mark. (5)	Mgmt. (2003–2005) (6)	N-Mark. (2003–2005) (7)
Size_{t-1}	−6.87***	−7.93***		−6.14***	−1.83***	−0.46	−1.85***
	(0.61)	(0.98)		(0.97)	(0.55)	(0.44)	(0.53)
Age_{t-1}	−2.98***	−10.50***	−0.53**	−1.43	−9.07***	−6.01***	−7.21***
	(1.01)	(1.69)	(0.26)	(1.64)	(1.21)	(1.19)	(1.41)
Co. Size_{t-1}	−1.16	0.53	−3.25*	7.87***	−7.33***	−4.40***	−8.37***
	(0.77)	(1.05)	(1.72)	(0.95)	(0.63)	(0.72)	(0.82)
# funds_{t-1}	−0.16*	−0.37***	−0.20	−0.71***	0.34***	−0.02	0.44***
	(0.09)	(0.12)	(0.33)	(0.12)	(0.07)	(0.07)	(0.08)
Turnover_{t-1}	−0.58	−2.64*	−2.05	−5.21***	2.58**	1.09	1.29
	(1.12)	(1.48)	(4.19)	(1.52)	(1.06)	(1.03)	(1.25)
σ_{t-1}	−5.51	3.82	41.91	−5.52	9.25	29.99*	38.60**
	(10.95)	(16.12)	(35.98)	(15.77)	(10.46)	(17.11)	(17.45)
Fair	4.77	5.03	−2.33	19.75***	−14.69***	−12.98***	−17.06***
	(4.50)	(5.81)	(21.28)	(4.52)	(4.39)	(3.06)	(4.38)
Good	−9.83**	−17.12***	−19.03	−1.94	−15.05***	−9.40***	−19.06***
	(4.40)	(5.93)	(19.17)	(4.73)	(4.19)	(3.17)	(4.22)
Excell.	−7.95	−13.74*	−38.63*	5.72	−19.33***	−10.69***	−19.52***
	(5.11)	(7.09)	(21.36)	(5.88)	(4.88)	(3.76)	(5.00)
$\widehat{\alpha}_t$		−50.70**	−104.83	−2.06	−48.57***	9.12	−17.50
		(22.70)	(116.06)	(15.41)	(15.08)	(24.14)	(40.89)
$\widehat{\alpha}_t$ Fair		28.15	35.53	−42.82**	71.05***	−3.01	67.03
		(24.52)	(116.67)	(18.42)	(16.81)	(27.19)	(44.59)
$\widehat{\alpha}_t$ Good		57.35**	157.18	1.90	54.74***	2.62	61.96
		(24.27)	(114.03)	(17.24)	(15.72)	(28.92)	(43.64)

(continued)

Table V—*Continued*

	TOC (1)	TOC (2)	TOC (3)	Mark. (4)	N-Mark. (5)	Mgmt. (2003–2005) (6)	N-Mark. (2003–2005) (7)
$\hat{\alpha}_t$ Excell.		61.63**	106.99	19.90	41.32**	−72.74**	−41.67
		(26.42)	(112.55)	(18.60)	(17.54)	(35.48)	(51.16)
S_t		−8.59	−7.71	6.01*	−14.91***	−6.08	−17.22***
		(5.37)	(18.35)	(3.52)	(4.08)	(3.98)	(4.37)
S_tFair		3.60	−12.43	−9.15**	12.99***	6.14	17.52***
		(5.94)	(19.30)	(4.37)	(4.36)	(4.19)	(4.69)
S_tGood		7.31	6.28	−3.09	10.56**	4.11	12.90***
		(5.70)	(18.05)	(4.00)	(4.23)	(4.33)	(4.58)
S_tExcell.		4.97	15.06	−12.21**	17.14***	10.31*	16.08***
		(6.50)	(19.28)	(4.81)	(4.93)	(5.43)	(5.74)
Obs.	7,767	3,933	545	3,933	3,971	1,734	1,734
Adj. R^2	58.58	67.39	68.64	65.73	46.59	39.51	39.40

Because as mentioned earlier, funds managed by the same management company frequently share the same board, we also reestimate the fee regression at the management company level by taking asset-weighted averages of the variables. Column (3) in Table V reports that the signs of the coefficients of interest are the same as in the previous column, although no coefficient (other than the intercept for funds with an Excellent grade) is significantly different from zero (a result that is not due to a reduced magnitude of the coefficients but to much larger standard errors).

Tufano and Sevick (1997) argue that the effect of fund governance should be most visible for management fees, since for these fees we can expect the conflict of interest between the management company and shareholders to be most severe. However, these authors find little difference in how board characteristics are related to marketing and nonmarketing fees. We replicate the analysis carried out in columns (1) and (2) with marketing and nonmarketing fees as dependent variables, although we report only the estimated coefficients of the full model in columns (4) and (5) of Table V, respectively. The results for marketing fees are mixed. Funds with Good and Excellent grades charge lower fees (although in column (4) the intercept for funds with an Excellent grade is marginally higher than that for funds with a Poor grade, the former are cheaper on average, other things equal) and have the highest coefficients on before-fee performance. However, the funds with a Fair grade—not those with a Poor grade—are the ones that charge the highest fees and have the lowest coefficient on before-fee performance. Therefore, there is no monotonicity at the bottom in the relation between board quality and marketing fees. Further, the differences between funds with a Poor grade and funds with a Good or Excellent grade are not statistically significant. Finally, the sensitivity coefficients follow no clear pattern: The lowest (and negative) coefficient corresponds to funds with the highest board quality grade.

The results for nonmarketing fees are more in line with those in columns (1) and (2). When we estimate the same specification as in column (1) (see the

Internet Appendix for the results), we find that other things equal, funds with a Poor grade charge the highest nonmarketing fees, and funds with Good and Excellent grades charge the lowest nonmarketing fees (which is the same ordering as the one implied by the intercepts in column (5)). The relation between performance sensitivity and nonmarketing fees becomes flatter as board quality improves: The sensitivity coefficient is -14.91 for funds with a Poor grade and the estimated differences imply that the same coefficient for funds with an Excellent grade is 2.23. However, further tests indicate that this coefficient is not significantly different from zero. Finally, funds with a Poor grade have a negative coefficient on performance. This coefficient is significantly lower than the coefficients for all other grades, although the coefficients for the higher grades are not increasing in board quality. Therefore, the results are consistent with the hypothesis that better-governed funds resist attempts by management companies to exploit investors' low performance sensitivity. The results also provide weaker evidence for the hypothesis that better board quality brings nonmarketing fees more in line with fund performance.

Since CRSP started providing data on management fees as of 2003, we can perform the analysis with management fees as the dependent variable for the 2003 to 2005 period. When we estimate a specification such as the one in column (1), we again find that the worst-governed funds charge the highest fees, as suggested by the ordering of the intercept coefficients reported in column (6). Further, the differences between the sensitivity coefficients are similar to those in column (5), although the statistical significance of the coefficients is marginal at best. However, the performance coefficients follow no clear pattern: The coefficient for the best-governed funds is not only negative but lower than all other coefficients, a result that is in stark contrast with those in columns (1) to (5). To check whether this result is due to a different behavior of management fees or to the different sample period, we reestimate the equation with nonmarketing fees, which are the closest approximation to actual management fees, as the dependent variable for the 2003 to 2005 period, and report the results in column (7) of Table V. Although the size and statistical significance of the performance coefficients vary, the ordering of these coefficients is the same in columns (6) and (7), showing that the best-governed funds have the lowest performance coefficient in both columns. We cannot explain this result. It may be due to an effort by the best-performing among the best-governed funds to lower their fees to compensate mutual fund investors for the decline in the stock markets during the 2000–2002 period, or to a response by these funds to the series of mutual fund scandals that emerged in 2003 and 2004. However, the short sample period, the noise inherent in our performance measure, and the fact that the only significant differences between the full sample and the 2003 to 2005 period have to do with the coefficients on performance suggest caution when interpreting the differences as indicating a structural change.

V. Conclusion

In this paper we show that there is a negative relation between funds' before-fee performance and the fees they charge to investors. Since this evidence is at odds with economic intuition, we subject it to a series of robustness tests, and find that it survives all of them.

Next, we propose two explanations for this anomalous result. According to the first explanation, the negative relation is the consequence of factors, which are omitted in univariate regressions, that are both positively correlated with returns and negatively correlated with funds' operating costs, and thus also with fund fees. According to the second explanation, in contrast, the negative relation is the result of funds that strategically set fees as a function of their past or expected performance. We consider three related rationales for this strategic behavior. The first, proposed by Christoffersen and Musto (2002), argues that funds with worse past performance have a pool of investors that are less sensitive to fund performance. Faced with an inelastic demand for their shares, underperforming funds optimally increase fees. The second explanation, proposed by Gil-Bazo and Ruiz-Verdú (2008), argues that funds with lower expected performance optimally set higher fees and target performance-insensitive investors, since these funds anticipate that they will not be able to compete with better-performing funds in the market for sophisticated investors. On the other hand, better-performing funds keep fees low because of competition among them for the money of performance-sensitive investors. The third explanation argues that funds with different expected performance choose different marketing strategies. Funds with low expected performance are marketed to performance-insensitive investors and have higher distribution costs, which translate into higher fees.

The empirical analysis finds support for all strategic-pricing explanations. Even though funds' operating costs are important determinants of fees, they do not explain away the negative relation between before-fee performance and fees. When we control for those cost determinants, we find that funds with lower expected before-fee performance and funds with less elastic demands charge higher marketing and nonmarketing fees. Therefore, it appears that mutual fund competition and regulation have not been sufficient to ensure that fees reflect the value that funds create for investors. However, we find some evidence that better fund governance may be associated with fees that are more in line with performance: Among the best-governed funds, worse performance need not mean higher fees.

Appendix: Estimation of Flow-to-Performance Sensitivity

We define annual net flow to fund i in year t, $Flow_{it}$, as the relative growth of the fund's total net assets (TNA) adjusted for returns net of expenses, R^n_{it}:[15]

$$Flow_{it} = \frac{TNA_{it} - TNA_{it-1}(1 + R^n_{it})}{TNA_{it-1}}. \tag{A1}$$

[15] Elton, Gruber, and Blake (2001) report a number of errors associated with mutual fund mergers and splits in the CRSP sample. Huang, Wei, and Yan (2007) argue that these errors could lead to extreme values of flows. We deal with this problem by eliminating the 1% of observations with the lowest and highest flows in each year.

Our model for fund flow determination builds on the main stylized facts that have emerged from prior studies on fund flows.[16] These studies show that flows of money to mutual funds are positively related to recent relative after-expense performance. In turn, the sensitivity of flows to performance has been shown to be higher for recent top performers, implying a convex flow-to-performance function, and lower for older funds. Further, the flow-performance curve becomes less convex as investor participation costs decrease (Huang, Wei, and Yan (2007)). Finally, research shows that flows depend on fund size, age, and expenses, total complex size, lagged flows, and total flows into funds with the same investment objective.

In addition to the variables considered in previous studies, we include the proxy for performance sensitivity proposed by Christoffersen and Musto (2002). These authors posit that funds that have experienced the largest outflows are left with the least performance-sensitive investors. Thus, these authors propose the following measure of fund attrition as a proxy for performance sensitivity:

$$Q/MAX_{it} = \frac{TNA_{it}}{MAX_{it}}, \tag{A2}$$

where TNA_{it} is fund i's total net asset value at the beginning of period t and MAX_{it} is the maximum total net asset value of fund i in the time-span up to period t.

We propose the following model for fund flow determination:

$$\begin{aligned} Flow_{it} = {} & a_{0t} + b_0 Perf_{it-1} + b_1 Perf_{it-1} rel\_age_{it-1} + b_2 Perf_{it-1} rel\_Q/MAX_{it-1} \\ & + a_M I_{M,it-1} + b_M Perf_{it-1} I_{M,it-1} + a_M^{PC} I_{M,it-1} PC_{it-1} \\ & + b_M^{PC} Perf_{it-1} I_{M,it-1} PC_{it-1} + a_H I_{H,it-1} + b_H Perf_{it-1} I_{H,it-1} \\ & + a_H^{PC} I_{H,it-1} PC_{it-1} + b_H^{PC} Perf_{it-1} I_{H,it-1} PC_{it-1} + \mathbf{c}'\mathbf{w}_{it-1} + \varphi_{it}, \end{aligned} \tag{A3}$$

where φ_{it} is a generic error term. The proxy for past performance, $Perf_{it-1}$, is the fund's four-factor alpha in year $t-1$, net of expenses, and in excess of the mean performance of all funds with the same investment objective in that year. The term $I_{M,it}(I_{H,it})$ is a dummy variable that equals one if $Perf_{it}$ is in the middle (top) third of all funds with the same investment objective in year t. We include these variables to allow for a convex relation between performance and flows. The variables rel_age_{it} and rel_Q/MAX_{it} are, respectively, the log of the fund's age in years and the fund's Q/MAX, in excess of the category's average in year t. The variable PC is a proxy for participation costs. We consider two of the proxies proposed by Huang, Wei, and Yan (2007): total assets managed by the company (in excess of the category's average in that year) and a dummy variable that equals one if there is another fund managed by the same management company with performance in the top 5% of its category, that is, a "star" fund.

[16] Studies on mutual fund flow determinants include Ippolito (1992), Gruber (1996), Chevalier and Ellison (1997), Sirri and Tufano (1998), Jain and Wu (2001), Barber, Odean, and Zheng (2005), Nanda, Wang, and Zheng (2005), and Huang, Wei, and Yan (2007).

The vector of lagged control variables, w_{it-1}, comprises: fund size and age; front- and back-end loads; 12b-1 fee; nonmarketing expenses; dummy variables for share classes; return volatility; log of total net asset value for all funds under the same management company; the fund's net flow; total flows of money to all funds with the same investment objective; and *Q*/MAX. Regressions also include year dummies.

The estimation results for equation (A3), which are available in the Internet Appendix, are consistent with results from previous studies: Flows are positively related to past (relative) performance; the flow-performance relation is convex; flow-to-performance sensitivity decreases with fund age; and convexity increases with participation costs. Consistent with Christoffersen and Musto's (2002) conjecture, *Q*/MAX is positively associated with flow-to-performance sensitivity.

Finally, we compute our measure of flow-to-performance sensitivity as the first derivative of conditional expected flow for performance, given the estimated coefficients from (A3):

$$S_{it} = \frac{\partial E_{t-1}(Flow_{it})}{\partial Perf_{it-1}} = \hat{b}_0 + \hat{b}_1 rel\_age_{it-1} + \hat{b}_2 rel\_Q/MAX_{it-1} + \hat{b}_M I_{M,it-1} + \hat{b}_M^{PC} I_{M,it-1} PC_{it-1} + \hat{b}_H I_{H,it-1} + \hat{b}_H^{PC} I_{H,it-1} PC_{it-1}, \quad \text{(A4)}$$

where $E_{t-1}(\cdot)$ denotes the expectation operator conditional on the information set at time $t-1$. Although we compute two measures of flow-to-performance sensitivity, each corresponding to a different proxy for participation costs, given the similarity of results, we report only those corresponding to complex size as a proxy for participation costs.

REFERENCES

Alexander, Gordon J., Jonathan D. Jones, and Peter J. Nigro, 1997, Mutual fund shareholders: Characteristics, investor knowledge, and sources of information, Office of the Comptroller of the Currency, E&PA Working paper 97–13.

Barber, Brad M., Terrance Odean, and Lu Zheng, 2005, Out of sight, out of mind: The effects of expenses on mutual fund flows, *Journal of Business* 78, 2095–2120.

Bergstresser, Daniel B., John M. R. Chalmers, and Peter Tufano, 2009, Assessing the costs and benefits of brokers in the mutual fund industry, *Review of Financial Studies*, forthcoming.

Berk, Jonathan B., and Richard C. Green, 2004, Mutual fund flows and performance in rational markets, *Journal of Political Economy* 112, 1269–1295.

Brown, Stephen J., and William N. Goetzmann, 1995, Performance persistence, *Journal of Finance* 50, 679–698.

Capon, Noel, Gavan J. Fitzimmons, and Russ A. Prince, 1996, An individual level analysis of the mutual fund investment decision, *Journal of Financial Services Research* 10, 59–82.

Carhart, Mark M., 1997, On persistence in mutual fund performance, *Journal of Finance* 52, 57–82.

Carhart, Mark M., Jennifer N. Carpenter, Anthony W. Lynch, and David K. Musto, 2002, Mutual fund survivorship, *Review of Financial Studies* 15, 1439–1463.

Chalmers, John M. R., Roger M. Edelen, and Gregory B. Kadlec, 2001, On the perils of security pricing by financial intermediaries: The wildcard option in transacting mutual fund shares, *Journal of Finance* 56, 2209–2236.

Chevalier, Judith A., and Glenn D. Ellison, 1997, Risk taking by mutual funds as a response to incentives, *Journal of Political Economy* 105, 1167–1200.

Chevalier, Judith A., and Glenn D. Ellison, 1999, Are some mutual fund managers better than others? Cross-sectional patterns in behavior and performance, *Journal of Finance* 54, 875–899.

Christoffersen, Susan E. K., Richard B. Evans, and David K. Musto, 2005, The economics of mutual-fund brokerage: Evidence from the cross-section of investment channels, Working paper, McGill University.

Christoffersen, Susan E. K., and David K. Musto, 2002, Demand curves and the pricing of money management, *Review of Financial Studies* 15, 1499–1524.

Cohen, Randolph B., Joshua D. Coval, and Ľuboš Pástor, 2005, Judging fund managers by the company they keep, *Journal of Finance* 60, 1057–1096.

Daniel, Kent, Mark Grinblatt, Sheridan Titman, and Russ Wermers, 1997, Measuring mutual fund performance with characteristic-based benchmarks, *Journal of Finance* 52, 1035–1058.

Deli, Daniel N., 2002, Mutual fund advisory contracts: An empirical investigation, *Journal of Finance* 57, 109–133.

Elton, Edwin J., Martin J. Gruber, and Christopher R. Blake, 2001, A first look at the accuracy of CRSP mutual fund database and a comparison of the CRSP and Morningstar mutual fund databases, *Journal of Finance* 56, 2415–2430.

Elton, Edwin J., Martin J. Gruber, and Christopher R. Blake, 2003, Incentive fees and mutual funds, *Journal of Finance* 58, 779–804.

Elton, Edwin J., Martin J. Gruber, and Jeffrey A. Busse, 2004, Are investors rational? Choices among index funds, *Journal of Finance* 59, 261–288.

Evans, Richard B., 2009, Mutual fund incubation, *Journal of Finance*, forthcoming.

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Fama, Eugene F., and James D. MacBeth, 1973, Risk, return and equilibrium: Empirical tests, *Journal of Political Economy* 81, 607–636.

Ferris, Stephen P., and Don M. Chance, 1987, The effect of 12b-1 plans on mutual fund expense ratios: A note, *Journal of Finance* 42, 1077–1082.

Ferris, Stephen P., and Xuemin S. Yan, 2007, Do independent directors and chairmen matter? The role of boards of directors in mutual fund governance, *Journal of Corporate Finance* 13, 392–420.

Ferson, Wayne E., and Rudi W. Schadt, 1996, Measuring fund strategy and performance in changing economic conditions, *Journal of Finance* 51, 425–461.

Gabaix, Xavier, and David Laibson, 2006, Shrouded attributes, consumer myopia, and information suppression in competitive markets, *Quarterly Journal of Economics* 121, 505–540.

Gaspar, José-Miguel, Massimo Massa, and Pedro Matos, 2006, Favoritism in mutual fund families? Evidence on strategic cross-fund subsidization, *Journal of Finance* 61, 73–104.

Gil-Bazo, Javier, and Pablo Ruiz-Verdú, 2008, When cheaper is better: Fee determination in the market for equity mutual funds, *Journal of Economic Behavior and Organization* 67, 871–885.

Golec, Joseph H., 2003, Regulation and the rise in asset-based mutual fund management fees, *Journal of Financial Research* 1, 19–30.

Gruber, Martin J., 1996, Another puzzle: The growth in actively managed mutual funds, *Journal of Finance* 52, 783–810.

Guedj, Ilan, and Jannette Papastaikoudi, 2005, Can mutual fund families affect the performance of their funds? Working paper, University of Texas at Austin.

Hadi, Ali S., 1994, A modification of a method for the detection of outliers in multivariate samples, *Journal of the Royal Statistical Society, Series B-Methodological* 56, 393–396.

Härdle, Wolfgang, 1990, *Applied Nonparametric Regression* (Cambridge University Press, Cambridge).

Huang, Jennifer, Kelsey D. Wei, and Hong Yan, 2007, Participation costs and the sensitivity of fund flows to past performance, *Journal of Finance* 62, 1273–1311.

Ippolito, Richard A., 1992, Consumer reaction to measures of poor quality: Evidence from the mutual fund industry, *Journal of Law and Economics* 35, 45–70.

Jain, Prem C., and Joanna S. Wu, 2001, Truth in mutual fund advertising: Evidence on future performance and flows, *Journal of Finance* 55, 937–958.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng, 2005, On the industry concentration of actively managed equity mutual funds, *Journal of Finance* 60, 1983–2011.

Kosowski, Robert, Allan Timmermann, Russ Wermers, and Halbert White, 2006, Can mutual fund "stars" really pick stocks? New evidence from a bootstrap analysis, *Journal of Finance* 61, 2551–2595.

Kuhnen, Camelia M., 2005, Dynamic contracting in the mutual fund industry, Working paper, Stanford University.

Latzko, David A., 1999, Economies of scale in mutual fund administration, *Journal of Financial Research* 22, 331–339.

Luo, Guo Y., 2002, Mutual fund fee-setting, market structure and mark-ups, *Economica* 69, 245–271.

Malhotra, D. K., and Robert W. McLeod, 1997, An empirical analysis of mutual fund expenses, *Journal of Financial Research* 20, 175–190.

Meschke, Felix, 2007, An empirical examination of mutual fund boards, Working paper, University of Minessota.

Morningstar Inc., 2006, Fact sheet: The Morningstar Stewardship Grade for funds. URL: http://quicktake.morningstar.com/DataDefs/StewGradeMethodology.pdf.

Nanda, Vikram, Z. Jay Wang, and Lu Zheng, 2005, The ABCs of mutual funds: On the introduction of multiple share classes, Working paper, University of Michigan.

Petersen, Mitchell A., 2009, Estimating standard errors in finance panel data sets: Comparing approaches, *Review of Financial Studies* 22, 435–480.

Sharpe, William F., 1966, Mutual fund performance, *Journal of Business* 39, 119–138.

Sirri, Erik R., and Peter Tufano, 1998, Costly search and mutual fund flows, *Journal of Finance* 53, 1589–1622.

Thompson, Samuel B., 2006, Simple formulas for standard errors that cluster by both firm and time, Working paper, Harvard University.

Tufano, Peter, Michael Quinn, and Ryan Taliaferro, 2006, Live prices and stale quantities: $T + 1$ accounting and mutual fund mispricing, Working paper, Harvard University.

Tufano, Peter, and Matthew Sevick, 1997, Board structure and fee-setting in the U.S. mutual fund industry, *Journal of Financial Economics* 46, 321–355.

Warner, Jerold B., and Joanna S. Wu, 2006, Why do mutual fund advisory contracts change? Fund versus family influences, Working paper, University of Rochester.

Wellman, Jay W., and Jian Zhou, 2008, Corporate governance and mutual fund performance: A first look at the Morningstar Stewardship Grades, Working paper, SUNY at Binghamton.

Wermers, Russ, 2000, Mutual fund performance: An empirical decomposition into stock-picking talent, style, transaction costs, and expenses, *Journal of Finance* 55, 1655–1703.

White, Halbert, 1980, A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity, *Econometrica* 48, 817–838.

50

AUTHOR:

Wesley R. Gray, Ph.D.
Executive Managing Member

ARE YOU TRYING TOO HARD?[1,2]

THE CASE FOR SYSTEMATIC DECISION-MAKING

EXECUTIVE SUMMARY

Everyone makes mistakes. It's part of what makes us human. Because humans understand their actions are sometimes flawed, it was perhaps inevitable that the field of psychology developed a rich body of academic literature to describe why it is that human beings often make poor decisions. Although insights from academia can be highly theoretical, our everyday life experiences corroborate many of these findings at a basic level: *"I know I shouldn't eat the McDonalds BigMac, but it tastes so good."* Because we recognize our frequent irrational urges, we often seek the judgment of experts, to avoid becoming our own worst enemy. We assume that experts, with years of experience in their particular fields, are better equipped and incentivized to make unbiased decisions. But is this assumption valid? A surprisingly robust, but neglected branch of academic literature, has studied the assumption that experts make unbiased decisions for over 60 years. The evidence tells a decidedly one-sided story: systematic decision-making, through the use of simple quantitative models with limited inputs, outperforms discretionary decisions made by experts. This essay summarizes research related to the "models versus experts" debate and highlights its application in the context of investment decision-making. Based on the evidence, investors should de-emphasize their reliance on discretionary experts, and should instead approach investment decisions with systematic models. To quote Paul Meehl, an eminent scholar in the field, *"There is no controversy in social science that shows such a large body of qualitatively diverse studies coming out so uniformly in the same direction as this one [models outperform experts]."*

*I THANK SEMINAR PARTICIPANTS AT THE UNIVERSITY OF AKRON, SEI INVESTMENTS, AND THE CLEVELAND CHAPTER OF THE AMERICAN ASSOCIATION FOR INDIVIDUAL INVESTORS FOR COMMENTS. I WANT TO THANK DAVE BABULAK, TOBY CARLISLE, TOM FRANK, RAIFE GIOVINAZZO, PAT MITCHELL, ANTHONY CAMBIERO, DOUG PUGLIESE, MARVIN KLINE, GABE KATES, AND EDDIE STERN FOR KEEN INSIGHTS. I ALSO THANK MY TEAM AT EMPIRITRAGE—DAVID FOULKE, JACK VOGEL, CARL KANNER, TAO WANG, YANG XU, AND TIAN YAO—FOR THEIR CRITICAL REVIEWS. ALL ERRORS ARE MY OWN.

---

[1] The title concept was inspired by Dean Williams 1981 keynote speech at Rockford College.

Source: http://turnkeyanalyst.com/wp-content/uploads/2013/02/Williams-Trying_too_Hard.pdf

[2] The inspiration for this piece is based on Toby and I's book, *Quantitative Value,* and James Montier's 2006 Dresner Kleinwort article, "Painting by Numbers: Ode to Quant."

ARE YOU TRYING TOO HARD?

SECTION 1: INTRODUCTION

"If you do fundamental trading, one morning you feel like a genius, the next day you feel like an idiot...by 1998 I decided we would go 100% models...we slavishly follow the model. You do whatever it [the model] says no matter how smart or dumb you think it is. And that turned out to be a wonderful business."

--Jim Simons, Founder, Renaissance Technologies[3]

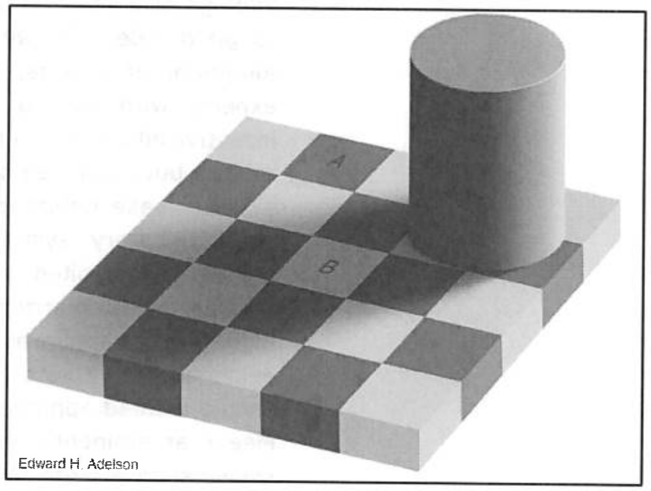
I should probably admit something up front: I once believed I was going be the next Warren Buffett. As a child, I raised animals and sold them at the county fair to make money. And with my growing savings came decisions—what to do with the money? To jumpstart my learning, my Grandmother gave me a copy of Benjamin Graham's *The Intelligent Investor,* which describes the philosophy of value-investing. I was 12 at the time and instead of being overwhelmingly appreciative, I was secretly depressed I didn't get a Nintendo. Nonetheless, I read the book and loved it. I was hooked on value-investing. Over the next 10 years I devoured books on value investing and eventually put my hard-earned "skills" to work, investing in value stocks and special situations.

Part of my investing education included matriculating in the finance PhD program at the University of Chicago. The first two years of the program were similar to drinking from a high-powered fire hose, which spewed sometimes unintelligible information and math equations from the leading scholars in finance. It was not always the most enjoyable experience. However, I persevered and met Professor Nick Barberis[4], who was researching the intersection between financial economics and psychology, a growing field that has since come to be known as "Behavioral Finance." I took Professor Barberis' PhD seminar and read over a hundred academic papers on behavioral finance. Although I wasn't sure how I could apply my new knowledge, I recognized that psychology was a powerful force in understanding financial economics.

Simultaneous with my exposure to behavioral finance, I was managing a small amount of money I had raised from my family and friends. I soon realized that the "irrational, emotionally involved, overconfident traders" Professor Barberis was referring to in his course weren't just theoretical investors dreamed up in the ivory tower—this crazy investor was me! I realized that no matter how many times I foolishly told myself that I was as smart as Warren Buffett, I would never actually be Buffett. I would always succumb to my innate cognitive biases. I guess sometimes it takes getting a PhD to realize you really don't know it all.

I also understood that I am not the only one capable of illogical thought—we all can succumb to bias. Figure 1 highlights this point.[5] Stare at box A and box B in the figure. If you are a human being you will identify that box A is darker than box B.

FIGURE 1



Then ask yourself:

"How much would I bet that A is darker than B?" $5? $20? $100?

We know how a human approaches this question, but how does a computer think about this question? A computer identifies the red-green-blue (RGB) values for a pixel in box A and the RGB values for a pixel in B. Next the computer tabulates the results: 120-120-120 for box A; 120-120-120 for box B. Finally, the computer compares the RGB values of the pixel in A and the pixel in B, identifies a match, and concludes that box A and box B are the exact same color. The results are clear to the computer.

Now, after taking into consideration the results from the computer algorithm, would you still consider A darker

3 http://video.mit.edu/watch/mathematics-common-sense-and-good-luck-my-life-and-careers-9644/, Accessed 2/10/2014

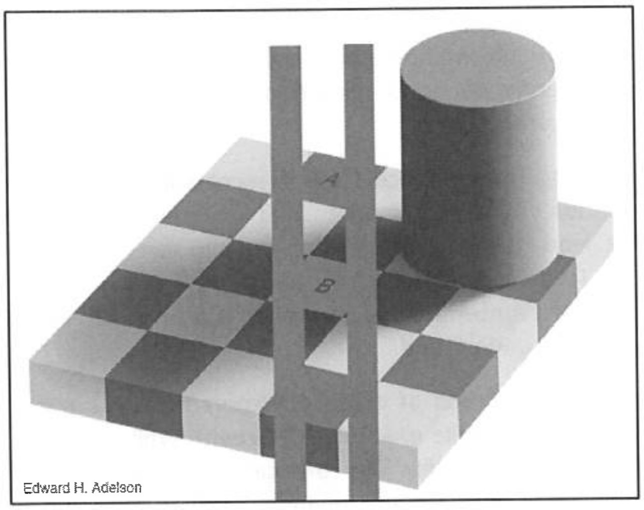
4 Dr. Barberis is now a professor at Yale University.

5 http://persci.mit.edu/gallery/checkershadow, accessed 2/10/2014.

han B? I don't know about you, but I still think A looks darker than B—call me crazy. (See Figure 2). But then that's what makes me human.

The sad reality is the computer is correct, and our perception is wrong. Our mind is being fooled by an illusion created by a vision scientist at MIT, Professor Ed Adelson. Dr. Adelson exploits local contrast between neighboring checkers, and the mind's perception of the pillar casting a shadow. The combination creates a powerful illusion that tricks every human mind. The human mind is, as succinctly stated by Duke psychology professor Dan Ariely, "Predictably irrational."

FIGURE 2



That may seem to be a strong statement. Perhaps the illusion above has convinced you that our minds may not be perfect in certain isolated settings. Or perhaps it has only persuaded you to believe that while a subset of the population may be flawed, you still possess a perfectly rational and logical mind. Don't be too sure, as a well-established body of academic literature in psychology demonstrates conclusively that humans are prone to poor decision making across a broad range of situations.

But are experts beyond the grip of cognitive bias? We often assume that professionals with years of experience and expertise in a particular field are better equipped and incentivized to make unbiased decisions. Unfortunately for experts, the academic evidence is emphatic: systematic decision-making, or models, outperform discretionary decision-making, or experts.

SECTION 2: ARE EXPERTS WORTHLESS?

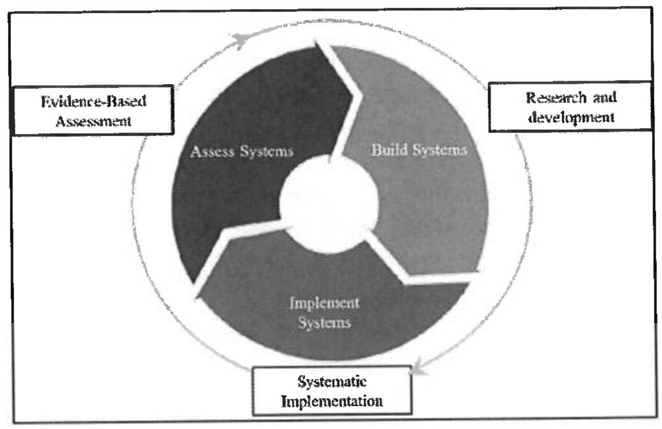
To be clear: I am not making the claim that human experts are worthless in the decision-making process. Experts are critical, but only for certain aspects of the decision-making process.

Students of decision-making break the decision-making process into three components (see Figure 3):

1. **Research and development**
2. **Implementation**
3. **Assessment**

I would argue that human experts are required for the first and third phases of a decision-making process, which are the research and development phase and the assessment phase, respectively. The crux of my argument is that human experts should not be involved in the second phase of decision-making, or the *implementation phase*.

FIGURE 3



During the research and development phase of decision making we build and test new ideas. In this phase, experts are *required* to create a sensible model. In the second phase—implementation—we should eliminate human involvement and rely on systematic execution. Finally, during the assessment phase of decision-making, we should once again rely on human experts to analyze and assess model performance to make improvements and incorporate lessons learned from the implementation phase.[6]

I look to the real world for insights into how this three-phase decision-making framework might be applied in practice. A great case study exists within the US Marine Corps, where I spent nearly four years as an officer deployed in a variety of combat situations. The USMC relies on "standard operating procedures," or SOPs, particularly when it places its Marines in harm's way. SOPs are developed according to the three-step process mentioned above, which is designed to establish the most robust, effective, and systematic decision-making

[6] Advancements in "machine learning" and related technologies might change this over time.

process possible. One example is the SOP for setting up a defensive position in a combat situation.[7] In the first phase of SOP development, experienced combat veterans and expert consultants review past data and lessons from the field to develop a set of rules that Marines will follow when establishing a defensive position. These rules are debated and agreed upon in an environment that emphasizes slow, deliberate, and critical thought. The current rules, or SOP, for a defensive position's priority of work is summarized by the acronym SAFE:

- **S**ecurity
- **A**utomatic Weapons on Avenues of Approach
- **F**ields of Fire
- **E**ntrenchment.

During the second phase—implementation phase—of SAFE, Marines in combat are directed to "follow the model," or adhere to the SOP. The last thing a Marine should do is disregard SOPs in the middle of a fire-fight, when the environment is chaotic and human decisions are most prone to error. Marines are trained from the beginning to avoid "comfort-based" decisions and to follow standard operating procedures. Of course, once the battle is over, Marines in the field will conduct a debrief and send this information back to the experts who can debate and assess in a calm environment whether the current SOP needs to be changed based on empirical experience gleaned from the field—the third phase. A key principle of this 3-step decision-making process is that discretionary experts are required to develop and assess, but execution is made systematic, so as to minimize human error. The Marines, like other critical decision-makers, want experts to develop and assess SOPs in a stable environment. However, the Marines want to implement SOPs ***systematically*** when the environment shifts from the war-gaming room to the live battlefield.

SECTION 3: THE EXPERT HYPOTHESIS

The so-called "expert's hypothesis," which asserts that experts can outperform models, is intuitive and tells a deceptively compelling story. For example, to most, it seems like common sense that a hedge fund manager with a Harvard MBA and 20 years of work experience at Goldman Sachs can beat a simple model that buys a basket of low P/E stocks. The logic behind this presumption is persuasive, as the expert would seem to possess a number of advantages over the model. The expert can arguably outperform the simple model for the following reasons:

- Experts have access to qualitative information.
- Experts have more data.
- Experts have intuition and experience.

Of course, there are other ways to support the argument that a human expert will beat a simple model, but most of these stories revolve around the same key points outlined above.

Three specious arguments underlie the expert's hypothesis:

- Qualitative information increases forecast accuracy.
- More information increases forecast accuracy.
- Experience and intuition enhance forecast accuracy.

Remarkably, the evidence I will present shows that soft information, more information, and experience/intuition do not lead to more accurate or reliable forecasts, but instead lead to poor decision-making. And because this result is so counterintuitive, it makes it that much more important to understand.

Among the hundreds of cases of expert forecasts gone awry, one high profile example is Meredith Whitney.[8] Ms. Whitney is famous for her prescient forecast of the banking crisis that reared its ugly head in late 2008. Public accounts of Ms. Whitney's predictions, widely observed and discussed during that time period, all suggested that Ms. Whitney was a "genius" after her remarkable call on Citibank's balance sheet blues.

Continued on next page...

[7] Marine Rifle Squad, MCRP 3-11.2, Ch 5.

[8] I do not mean to single out Meredith Whitney. The same point can be made with just about any analyst who has shown up on CBNC and expressed a confident and detailed opinion on a forecast.

FIGURE 4

"50-100 sizable municipal defaults" –Whitney, 2010

"There were just 5 defaults" –WSJ, 2013

But Ms. Whitney didn't stop there. She outlined her gloomy forecast for the municipal bond market on a December 2010 segment of the prime-time CBS news program, *60 Minutes*. Ms. Whitney predicted there would be "50 to 100 sizable defaults." She forcefully reiterated her prediction at the Spring 2012 Grant's Interest Rate Observer Conference, where observed firsthand the emotional conviction Ms. Whitney felt for her bold prediction.

However, Ms. Whitney's powers of prediction were fleeting. In an article published in September 2012, the *Wall Street Journal* published a stinging article entitled, "Meredith Whitney Blew a Call—And then Some." The piece was quick to point out that that "there were just 5 defaults" in the municipal market.[9] (See Figure 4).

Ms. Whitney was off by a factor of 10.

Whitney's missed call embodies the assumptions underlying the expert hypothesis. She was a well-known expert with access to the best qualitative and quantitative data available. That, coupled with her well-known previous experience and astute intuition, made her story compelling to the media and other experts alike. Many believed that Whitney *had* to be right. Whitney, like everyone else, also thought she *had* to be right. She had access to important people in local and state governments who provided her with privileged "soft" information; she studied thousands of pages of municipal bond term-sheets and macroeconomic research reports; and her recent experience making the call on the Financial Crisis crystallized in her own mind that she could trust her "gut." Unfortunately, the potent combination of realized success and intense effort, gives human experts the "illusion of skill,"[10] which translates into overconfidence and a failure to appreciate randomness. Mark Twain's quip summarizes the problem: "It ain't what you don't know that gets you into trouble. It's what you know for sure that just ain't so." In other words, our highest conviction decisions are apt to cause us the most problems.

SECTION 4: THE EVIDENCE ON MODELS VERSUS EXPERTS[11]

To substantiate the argument that the expert hypothesis is false, I stand on the shoulders of academic researchers who have studied this hypothesis for over half a century.

To give readers a flavor for how academic research has studied the relative performance of models versus experts, I introduce a study on parole recidivism predictions.[12] The study was facilitated by a partnership between a group of academic researchers and the Pennsylvania Board of Probation and Parole to identify ways to make the parole process more accurate, fair, and cost-effective. Professor Carroll and his team set up the experiment as follows:

- **Experts:** Collect information to make parole board decisions based on quantitative and qualitative information used in the decision-making process. This included interviews with the parolees, interviews with people known by the parolee, information on past criminal history, demographics, and so forth.
- **Models:** Create a simple predictive model of parolee recidivism based on predictive factors. The baseline model consists of 3 elements: offense type, number of past convictions, and number of prison rule violations.

The researchers compare the performance of the experts against simple models in out-of -sample tests. The

---

[9] http://online.wsj.com/news/articles/SB10000872396390444549204578021380172883800: accessed, 2/10/2014.

[10] Kahneman, D. *Thinking, Fast and Slow*: New York, Macmillan, 2012, p. 212.

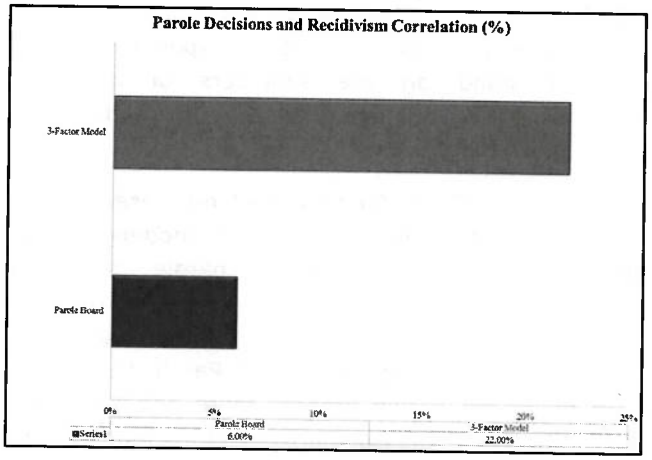
[11] In the jargon of academia, the term for experts is "clinician" and the term for models is "actuarial process." Instead of using "clinical versus actuarial," I use "models versus experts" to facilitate understanding within our chosen context.

[12] Carroll, J., Wiener, R., Galegher, J, and J. Alibrio, 1982, "Evaluation, Diagnosis, and Prediction in Parole Decision Making," *Law, and Society* Review 17, p. 199-228.

summary finding is that a simple model, while far from perfect, is over three times more effective than experts at forecasting recidivism (See Figure 5).

The research results were not lost on practitioners in the real world. A 2013 *Wall Street Journal* article, "State Parole Boards Use Software to Decide Which Inmates to Release," highlights how algorithmic parole decisions are now the norm, and no longer the exception.[13]

FIGURE 5



WHAT IF EXPERTS HAVE THE MODEL?

The evidence above suggests that models beat experts in the context of parole recidivism. Subsequent research came to similar conclusions across a variety of contexts pitting man versus machine. Perhaps more importantly, the research on models versus experts inspired scholars to tackle another question:

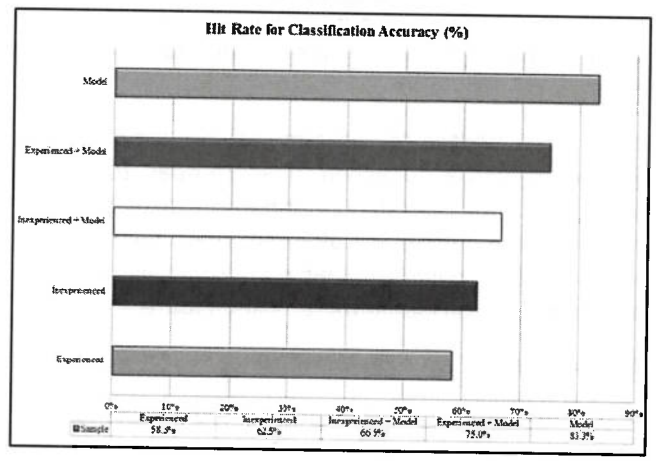
How do experts perform if they are given the results of the model?

D. Leli and S. Filskov explore this question in their 1984 study, "Clinical Detection of Intellectual Deterioration Associated with Brain Damage."[14] The study's premise is simple. First, place experienced psychologists and a simple prediction algorithm head-to-head in a horse race. Next, see who can more accurately classify the extent of a patient's brain impairment, based on W-B protocol statistics.[15] The model utilizes a systematic approach based on a statistical model of prior data; meanwhile, the humans can utilize their vast experience and intuition based on years of experience.

The results from the study are striking. The simple quantitative model has a classification accuracy ratio of 83.3%, much higher than the experienced clinicians, who had a success rate of only 58.3%. Interestingly, the inexperienced clinicians were slightly better at 62.5%. The model (and the novices) clearly beat the experts.

But the researchers took their analysis one step further. They wanted to explore what would happen when the experts were armed with the powerful prediction model. A natural hypothesis is that experts, equipped with the model, would outperform the stand-alone model. In other words, models might represent a floor on performance, to which the experts could add performance.

FIGURE 6



In follow-on tests, the researchers gave the experts the output of the model and disclosed that the model has "previously demonstrated high predictive validity in identifying the presence or absence of intellectual deterioration associated with brain damage." Using the model, experienced clinicians significantly improved their accuracy ratio from 58.3% to 75% and the inexperienced clinicians moved from 62.5% to 66.5%. Nonetheless, the experts were still unable to outperform the stand-alone model, which had already established the gold standard 83.3% success rate. This study suggests that models don't represent a floor on performance; rather, models reflect a ceiling on

---

13 http://online.wsj.com/news/articles/SB10001424052702304626104579121251595240852. Accessed 2/10/2014.

14 Leli, D., and S. Filskov, 1984, "Clinical Detection of Intellectual Deterioration Associated with Brain Damage," *Journal of Clinical Psychology* 40, p. 1435-1441.

15 W-B protocol includes information on age, sex, education, intelligence tests, and so forth.

performance, from which the experts detract. The "secret sauce" of human judgment ruins the beautiful simplicity of a calculation (See Figure 6).

BUT QUANTAMENTAL IS DIFFERENT, RIGHT?

Investors consider themselves to be unique, in the sense that there is a belief that a human expert armed with a model can generate outsized returns in financial markets, or in other words, financial models represent a floor on performance, not a ceiling. Carson Boneck, Global Head of Investment Management for S&P Capital IQ, stated the following in May 2013:

> *"We think quantamental is going to be a big theme driving our client portfolios and our own product strategy."*[16]

Mr. Boneck tells a great story: A discretionary stock-picking expert armed with a powerful model may create a powerful combination. But investment clients should ask if there is any evidence that a quantamental approach—a process that involves the use of models to screen for promising stocks, but overlays a human element to make the final investment decision—actually adds value. One might argue that the experts in the Leli and Filskov (1984) study were sub-par and perhaps the study design was flawed. Or perhaps these results are only relevant to the field of brain research. Expert stock pickers with years of experience in the investment management business, by contrast, may have access to superior fundamental research tools and can develop a more pronounced qualitative information edge. Stock pickers can't possibly be beaten by simple models, can they? As it turns out, we have a reasonable real-world laboratory that provides insight into this question.

Joel Greenblatt, famous for his bestselling books *You Can Be a Stock Market Genius* and *The Little Book that Beats the Market,* stumbled into a natural experiment. Joel's firm, Formula Investing, utilizes a simple algorithm that buys firms that rank high on an average of their cheapness and their quality ranking. The firm offers investors separately managed accounts (SMAs) and investors have a choice in how to invest:

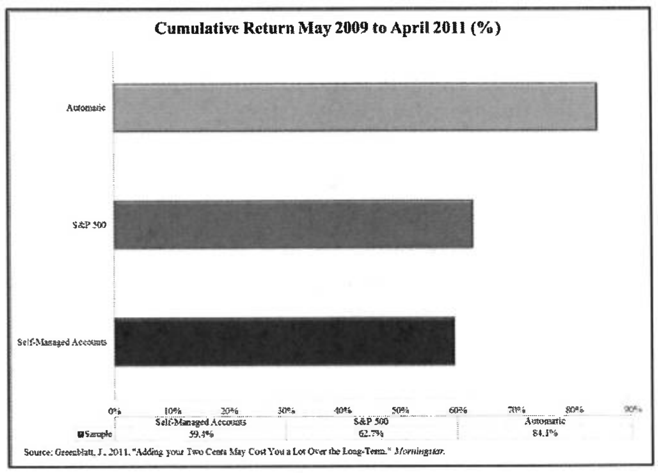
1) Invest using only the model's output.
2) Receive the model's output, but use discretion to identify stocks held in the portfolio.

[16] http://www.youtube.com/watch?v=tThxb_eFUTo. Accessed 2/10/2014.

Joel collected data on all their SMAs from May 2009 through April 2011 and tabulated the results. He wanted to know if adding discretion to the investment process would improve results. (See Figure 7).

The automatic accounts earned a total return of 84.1%, besting the S&P 500 Index's 62.7% mark by over twenty percentage points. The self-managed accounts, which allowed the clients to pick and choose from the model's output at their discretion, earned a respectable 59.4%. However, the 59.4% figure was worse than the passive benchmark, and much worse than the account performance for the automatic accounts.[17] This takeaway from this study is similar to the brain impairment research by Leli and Filskov: *Models represent a ceiling on performance, not a floor.*

FIGURE 7



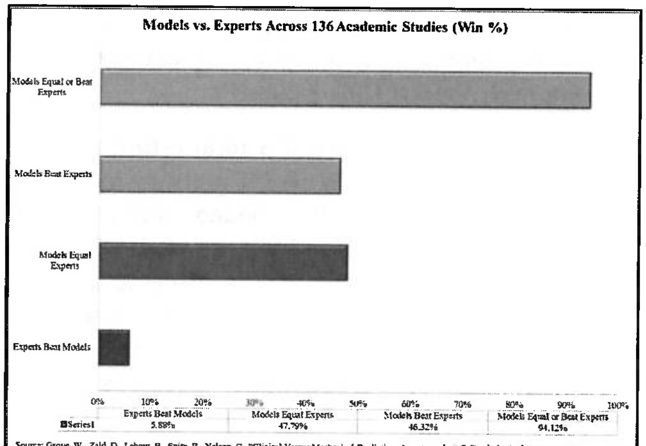
A STUDY OF ALL THE STUDIES

Thus far I've presented a few formal studies and an ad-hoc study of investor behavior. In order to make a more convincing case that models beat experts, I require more analysis. Luckily, one doesn't have to look far for additional evidence. Professors William Grove, David Zald, Boyd Lebow, Beth Snitz, and Chad Nelson have performed a meta-analysis—or a study of studies—on 136 published studies that analyze the accuracy of "actuarial" (i.e., computers/models) vs. "clinical" (i.e., human experts) judgment.[18]

[17] Greenblatt, J., 2011, "Adding your Two Cents May Cost You a Lot Over the Long-Term," Morningstar.

[18] Grove, W., Zald, D., Lebow, B., and B. Nelson, 2000, "Clinical Versus Mechanical Prediction: A Meta-Analysis," *Psychological Assessment* 12, p. 19-30.

FIGURE 8



Source: Grove, W., Zald, D., Lebow, B., Snitz, B., Nelson, C., "Clinical Versus Mechanical Prediction: A meta-analysis." *Psychological Assessment* 12 (2000), 19-30.

The studies examined by Grove et al. include forecast accuracy estimates that cut across many professions. The studies included examples from the following fields:

- College academic performance
- Magazine advertising sales
- Success in military training
- Diagnosis of appendicitis
- Business failure
- Suicide attempts
- University admissions
- Marital satisfaction
- Wine quality

The study's results are stunning: Models beat experts 46% of the time; models equal or beat experts 94% of the time; and experts beat models a mere 6% of the time (See Figure 8). The empirical evidence that systematic decision processes meet or exceed discretionary decision making would seem to be overwhelming. The executive summary of the study says it best:

> *Superiority for mechanical-prediction techniques was consistent, regardless of the judgment task, type of judges, judges' amounts of experience, or the types of data being combined. --Grove et al.*

The empirical evidence on the horse race between model-driven decisions and discretionary decision-making is clear, maybe even a slam dunk, but the implications are unsettling. How is it possible that a simple algorithm can consistently beat expert opinion? The answer to this conundrum lies with cognitive bias.

SECTION 5: WHY EXPERTS FAIL TO BEAT MODELS

Daniel Kahneman's magnum opus, *Thinking, Fast and Slow*,[19] describes a less-than-perfect view of human decision-making. The thesis of the book is that humans are driven by two modes of thinking: System 1 and System 2.

- **System 1:** Decisions are instinctual and heuristic-based.
- **System 2:** Processes are calculated and analytical.

System 1 thinking, while imperfect, is speedy and highly efficient. For example, if Joe is facing the threat of a large tiger charging him at full speed, his System 1 thinking will trigger Joe to turn around and sprint for the nearest tree, and ask questions later. As an alternative, Joe's System 2 thinking will calculate the speed of the tiger's approach and assess his situation. Joe will examine his options and realize that he has a loaded revolver that can take the tiger down in an instant. On average, if Joe immediately sprints to the tree he may get lucky and outrun the tiger; on the other hand, if Joe pauses and calculates his best option, which is blowing the tiger away with his revolver, Joe's tactical pause may end with Joe trying to remove a 500 pound meat-eating monster from his jugular vein.

Joe's hypothetical situation highlights why evolution has created System 1; on average, running for the tree is a life-saving decision when faced with a high-stress situation where survival is on the line. The issue with System 1 is that its heuristic-based mechanisms often lead to systematic bias: Joe will always run, even when the better decision may be to shoot.

System 1 certainly served its purpose when humans were faced with life and death situations in the jungles, but in modern day life, where decisions in chaotic work environments may have limited physical consequences, the benefits of immediate decisions rarely outweigh the costs of flawed decision-making. For example, the necessity of avoiding System 1 and relying on System 2 in the context of financial markets is of utmost importance.

I highlight below three core reasons why human experts making discretionary decisions underperform systematic decisions facilitated via simple models, with each discussed in more detail:

---

[19] Kahneman, D. *Thinking, Fast and Slow*: New York, Macmillan, 2012.

ARE YOU TRYING TOO HARD?

1. Same facts; different decisions

Humans, unlike models, can take the same set of facts and arrive at different conclusions. This can happen for a variety of reasons, but a lack of human consistency is often attributed to anchoring bias, framing effects, availability bias, or something as simple as hunger and fatigue. A computer suffers from none of these ailments—same input, same output.

2. Story-based, not evidence-based decisions

Humans suffer from a proclivity to believe in stories, or explanations that fit a fact pattern, but they don't bother to fully consider the empirical evidence. For example, consider the following statement:

Linda is thirty-one years old, single, outspoken, and very bright. She majored in philosophy. As a student, she was deeply concerned with issues of discrimination and social justice, and also participated in antinuclear demonstrations.

Is it more likely that Linda is a bank teller or that Linda is a bank teller and is active in the feminist movement? Our gut instinct is to think that it is more likely that Linda is a feminist bank teller, since the concepts of discrimination and social justice activate our stored memories that are traditionally associated with feminism. But this line of reasoning is incorrect, as it ignores the statistical reality.[20] System 1's love for a coherent story has led us to make a poor judgment, which is divorced from the true underlying probabilities. An empirical-based decision would consider the fact that the bank teller population is much larger than the feminist bank teller population and immediately understand that it is statistically more likely that Linda is a bank teller. We have been led astray by our urge to create a story that in our gut seems to describe the evidence.

3. Overconfidence

Humans are consistently overconfident. Overconfidence can be driven by cognitive errors such as hindsight bias—believing past events were more predictable ex-ante than they actually were—and self-attribution bias—attributing good outcomes to skill and poor outcomes to bad luck. Systematic decisions limit these problems. Models don't get emotionally involved and don't have an ego. Therefore, they are unable to get overconfident or overoptimistic—they simply execute based on the facts.

[20] Assume there are 200 females, 100 female bank tellers, and 50 female feminists in the world. It is more likely that Linda is a bank teller (100/200=50%) because the subset of bank tellers that are also feminist (best case is all feminist are bank tellers implies 50/200=25%) is much smaller than the population of bank tellers as a group (100).

SECTION 5A: SAME FACTS; DIFFERENT DECISIONS

Anchoring

Stimuli from the environment attack human discretionary decisions at a subconscious level. Often we don't even know we are vulnerable. One important example is anchoring. Broadly defined, anchoring describes our tendency to rely too heavily, or "anchor," on irrelevant information when making decisions.

An example comes from research by Professors Simonson and Drolet who study how consumer behavior is affected by irrelevant anchors.[21] The researchers ask buyers to assess their willingness to pay for a variety of products, including a *Black & Decker Cooltouch Two-Slice Toaster*. The researchers play a trick on some of the buyers along the way. The selected buyers are asked to write down the last 2 digits of their social security number prior to asking the question about willingness to pay. The anchoring hypothesis predicts that buyers with higher SSN values will be willing to pay a higher amount and those with lower SSN will be willing to pay a lower amount.

Remarkably, the value for the last 2 digits of one's social security number *actually influences the buyer's willingness to pay*. Buyers with SSNs above 50 report a willingness to pay of $32.50, whereas, buyers with SSNs less than 50 report $25. Those who are not affected by the SSN anchor report $30. The researchers repeat this experiment on different consumer products such as phones, backpacks, and radio headphones and find similar results. The evidence from this study—and the many other studies like it—document that anchoring effects have a powerful influence on our decisions.

Continued on next page...

[21] Simonson, I., and A. Drolet, 2004, Anchoring Effects on Consumers' Willingness-to-Pay and Willingness-to-Accept, *Journal of Consumer Research* 31, p. 681-690

ARE YOU TRYING TOO HARD?

FIGURE 9

AN EXAMPLE OF THE WTP WITH SSN ANCHOR TASK (STUDY 1)

TOASTER

Description: *Black & Decker Cooltouch Two-Slice Toaster.* Has super wide slots for extra thick bread, bagels and English muffins. Exterior stays cool to the touch. Removable crumb tray for easy cleaning.

First, please enter the last two digits of your social security number (SSN): ______

1. Now assume that the last two digits of your SSN are a price in dollars. Would you be willing to pay this price for this toaster (circle one)?

YES NO

2. What is the highest price you would be willing to pay for this toaster? $_______

FIGURE10

	WTP with SSN	
Last two digits of SSN	<50	>50
Product		
Toaster	$25	$32.5
Phone	50	70
Backpack	25	30
Radio Headphone	20	30
Average	*$30*	*$41*

But how might anchoring affect a professional stock-picking portfolio manager? Imagine the manager is conducting a discounted cash flow analysis that requires a 10-year revenue growth forecast. There are 2 scenarios:

1) The manager's secretary walks in and mentions that his prior meeting with a new client has been moved to the 5th of January
2) The secretary walks in and mentions that his prior meeting with a new client has been moved to the 30th of January.

The only difference in the 2 scenarios is the mention of "5th" and "30th." Is the manager going to input the same revenue growth projection in both these scenarios? If Professor Kahneman's description of the strength of the anchoring bias is true, it is likely he will enter a higher growth rate number in the "30th" scenario:

> *"[Anchoring is] one of the most reliable and robust results of experimental psychology."*[22]

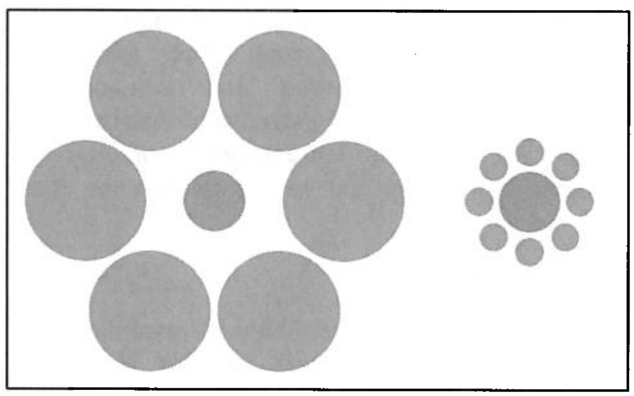
Research on anchoring suggests that the portfolio manager will enter different revenue growth numbers into his model in both of these scenarios. The scariest part is that the manager won't even know this occurring, because anchoring effects are influencing his decision-making process at the sub-conscious level.

Framing

Framing is another bias that creeps into our decisions without our knowing we are being influenced. The bias occurs because different ways of presenting information can evoke different emotions, which then leads to different decisions.

Consider the 2 orange circles in figure 11.[23] Even though the orange circles are the same size, it appears that the circle on the right is larger than the one on the left because of framing. The right circle is surrounded by smaller circles, causing our brains to perceive the right circle as being large when compared with the surrounding smaller circles. However, the left circle is surrounded by larger circles, leading our brain to interpret it as being relatively smaller than the right circle. Go ahead, stare at the circles as long as you'd like. Our brains are programmed to perceive the two circles differently, based on the context, or frames, even though they are exactly the same size. Wild!

FIGURE 11



[22] Kahneman, D. *Thinking, Fast and Slow*: New York, Macmillan, 2012, p. 119.

[23] Ariely, D., *Predictably Irrational*: New York, Harper, 2010.

ARE YOU TRYING TOO HARD?

Amos Tversky and his colleagues study framing by changing how questions are asked.[24] They find that the same information can be conveyed using two different frames and people will respond in completely different ways, depending on which frame is used.

I highlight a few representative examples below:

Do you prefer ground beef that is...

- *75% lean?*
- *25% fat?*

Who would ever want "fat" beef? Most people will choose ground beef that is 75% lean over 25% fat, without recognizing that the questions are exactly the same.

Another example:

Do you prefer a medication that has a ...

- *90% chance of keeping you alive?*
- *10% chance of killing you?*

Our brains immediately think "Staying alive or dying?—that's an easy question. I choose alive." Of course, we have to strain a bit to realize that the two propositions are exactly the same.

One doesn't have to think too hard to see how a financial advisor with training in psychology could influence her customers. Consider a financial advisor who tells her client the following:

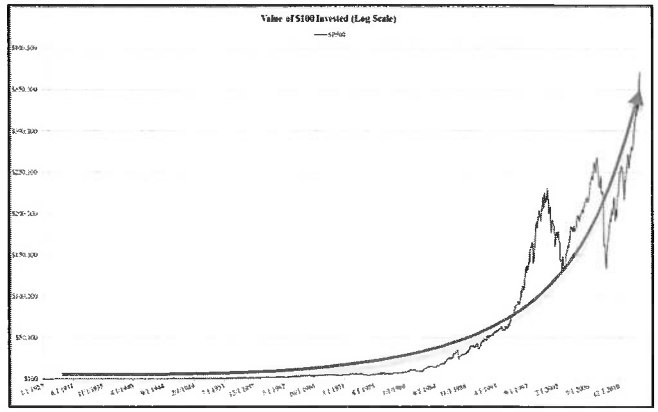
"Stocks are better for the long run because they earned 9.91% a year from 1927 through 2013."

Take that same financial advisor, but have her frame the information a bit differently:

"Stocks are better for the long run because they grew $100 into $371,452 from 1927 through 2013."

This same advisor frames the information in a chart:

FIGURE 12



Although the advisor technically offered the exact same information (stocks grew at 9.91% from 1927 to 2013), the framing of the second statement alongside a fancy chart will encourage the client to allocate more to stocks. The thought of turning $100 into $371,452 is much more appealing at first glance than earning a "measly" 9.91% a year, which is an abstraction with less immediately perceived value.

Availability Bias

Availability bias is an artifact of System 1, which causes our mind to overemphasize the importance of recent or easily recalled information. An applied example: Imagine someone asks you whether there are more English words that begin with a "k" or have "k" as the third letter? Your mind is slowly churning: *kid, kiss, key...*

Found any words with "k" as the third letter?

Probably not—it's difficult to recall these words. By contrast, words starting with a "k" spring effortlessly to mind. Naturally, words that start with a "k" must therefore be more prevalent in the English alphabet, right? Wrong. There are three times as many words that have "k" in the third position in the English language than there are words beginning with "k."[25]

[24] McNeil, Pauker, Sox Jr., and Amos Tversky, "On the elicitation of preferences for alternatives therapies," *New England Journal of Medicine* 306 (1982): 1259-62.

[25] A. Tversky and D. Kahneman, 1973, Availability: A heuristic for judging frequency and probability, *Cognitive Psychology* 5, p. 207-233.

FIGURE 13

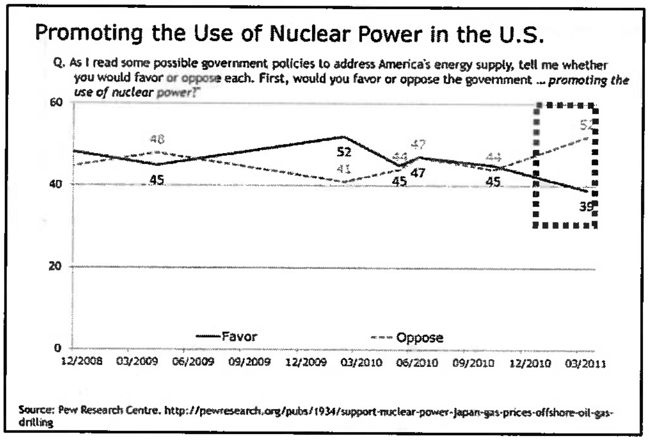
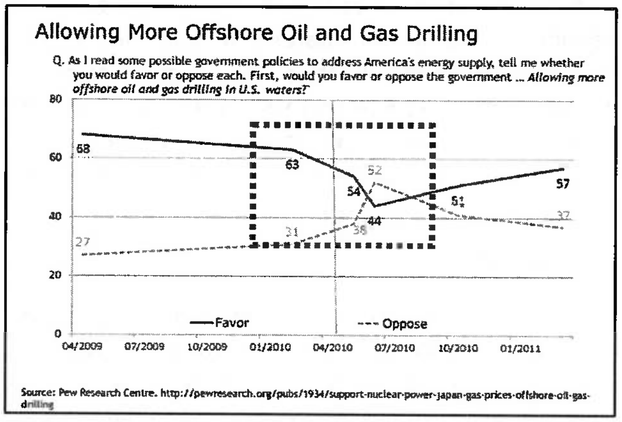


FIGURE 14



We see availability bias in a number of contexts when there is an "availability shock" in the news media. Two recent examples include the Fukushima Nuclear Disaster in Japan (see Figure 13) and the BP oil spill in the Gulf of Mexico (see Figure 14).

Pew Research Center conducts periodic sentiment polls on the use of nuclear power and oil and gas drilling in the United States. The polls clearly show how recent events—especially those spectacularly displayed in the media—can change implicit assessments of disaster probabilities. For example, sentiment on the use of nuclear power in the U.S. shifts from an even split in June 2010 to 52% opposed and 39% in favor in March 2011. Similarly, in March 2010, 62% favored and 31% opposed offshore oil and gas drilling. However, after the BP oil spill disaster in the Gulf of Mexico, the June 2010 poll numbers are 52% opposed and 44% in favor of offshore drilling.

Perhaps you don't believe this would affect investors? All one needs to do is read the Franklin Templeton Annual Global Investment Sentiment Survey, which asked investors—after the fact—to estimate the S&P 500 index performance for 2009 and 2010.[26] 66% of investors believed the S&P 500 was down or flat in 2009, when the S&P was up 26.5%; 49% thought the market was down or flat in 2010, when the S&P was up 15.1%. The massive drawdown associated with the 2008 Financial Crisis obviously left a stinging—and available—impression on market participants.

Physical State

> *"How were you feeling when you got out of bed thirteen years ago, when you're looking at historical simulations? Did you like what the model said, or did you not like what the model said? It's a hard thing to back-test."*
>
> *--Jim Simons, CEO, Renaissance Technologies, LLC*

Any coffee drinkers out there? If I don't have my coffee in the morning I feel sluggish and my head starts pounding. If you were to lay out a set of financials for a firm and demanded an earnings forecast, I guarantee you that my answer will be highly dependent on my coffee consumption that morning. Physical state, while often overlooked by those discussing behavioral finance, is probably the most intuitive and compelling reason why a human expert can have the same set of facts, and yet come to different conclusions.

An interesting empirical study highlights the power of basic biological impulses (circadian rhythm) on the human mind. Bodenhausen conducts a study highlighting varying degrees of discrimination exhibited by individuals who self-identify as either "morning types" or "evening types."[27] Each individual is asked, at different times of the day, to state his opinion on the guiltiness of a suspect associated with an alleged crime. The descriptions of the suspects are purposely stereotyped in a way that should appeal to innate discrimination, thereby triggering System 1 heuristic decision-making.

The author tabulates the results for the perceived guilt by time of day for all participants in the study. Figure 15 breaks out the results for the 9:00am and 8:00pm surveys for both morning and evening types. Morning types were more likely to give a suspect the benefit of

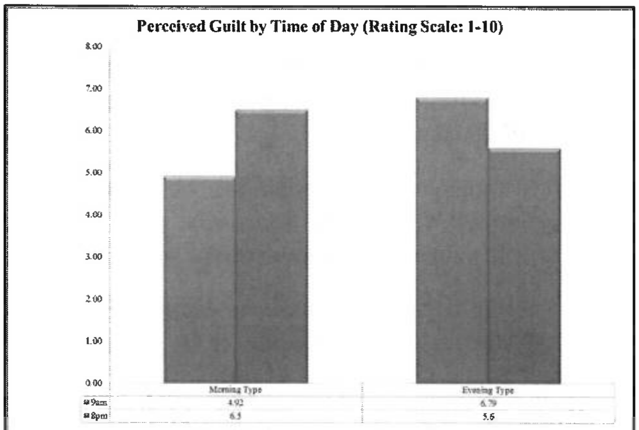
---

[26] The Franklin Templeton 2010/2011 Global Investor Sentiment Survey, https://www.franklintempleton.com/investorsentiment, Accessed 2/10/2014.

[27] Bodenhausen, G., 1990, Stereotypes as Judgmental Heuristics: Evidence of Circadian Variations in Discrimination, *Psychological Science* 1, p. 319-322.

the doubt when they were feeling bright and chipper early in the day, but much more likely to view the suspects as guilty when asked later in the day when their minds were wandering. Evening types exhibited the same pattern, but in reverse. Suspects were more likely to be considered guilty by the evening types in the morning, when they were presumably grumpier and less aware, but these same individuals were more lenient on suspects when asked in the evening, during a time when they were feeling awake.

FIGURE 15



Source: Bodenhausen G, 1990, Stereotypes as Judgemental Heuristics: Evidence of Circadian Variations in Discrimination, *Psychological Science* 1, p. 319-322.

The authors conclude that circadian rhythms that govern daily wakefulness patterns can affect decision-making. Is it fair or sensible that circadian pattern can affect the treatment of suspects on trial? Not really, but trial defense lawyers picking potential jurors for a morning trial should focus on identifying morning types who've had their 24 oz. Starbucks coffee!

SECTION 5B: STORY-BASED, NOT EVIDENCE-BASED DECISIONS

My daughter asks, "How did we get these presents under the Christmas tree, daddy?" I respond, "Oh, Santa dropped them off." She retorts, "How did he bring them here?" I reply, "On his sleigh guided by his reindeers." My daughter comes back, "Oh, yeah, that makes sense. He even ate the cookies we left by the fireplace and his reindeers ate the carrots we left outside."

Human beings have strong preferences for coherent stories and often build powerful narratives to help interpret complex situations. In my daughter's case, the impossible physics of the Santa story will gradually break down the story over time, but the powerful Santa Claus narrative will likely extend well beyond what many would consider its "rational life." The Santa Claus story is one that appeals to many young children[28], but it is not just they who suffer from a strong belief in stories, adults are susceptible as well.

Skinner's Pigeons

The foundation for our persistent belief in stories, in spite of evidence suggesting a story is literally unbelievable, has perplexed researchers for many years. The behavioral psychologist B.F. Skinner and several colleagues demonstrated that our innate need for superstition is deeply ingrained in our primal brains.[29] To make the point, Skinner studied one of the more powerful brains in the animal kingdom—the pigeon (see Figure 16[30]).

FIGURE 16

Skinner put hungry pigeons in a cage and dispensed food pellets to them every 5 seconds. Now, pigeons will naturally wander around any space looking for food, and will do so in predictably pigeon-like ways. One pigeon might step to the left and then step to the right; another pigeon might jump, land, and then jump again. Following these random movements a food pellet will inevitably appear, consistent with the five second release pattern. After a few rounds of engaging in the same random activities and earning a series of food pellets, the pigeons develop an internal story that their deliberate actions in the cage are *causing* food pellets to pop out of the feeder.

[28] In my case, I believed in Santa Claus until the age of 14—and still do at times. Physics be damned!

[29] B.F. Skinner, 1948, Superstition in the Pigeon, *Journal of Experimental Psychology* 38, p. 168-172.

[30] Image source unknown.

ARE YOU TRYING TOO HARD?

Amazingly, once a pigeon establishes such a superstition, it is exceedingly difficult to train the pigeon out of the story. Skinner attempts to give the pigeons evidence that their superstition is worthless, but the pigeons continue with their story-based ways. Evidence has a hard time entering the decision-making process once a behavior has been established.

Pigeons aren't the only animals suffering from "story bias." My uncle is convinced that a Dallas Cowboys victory during the Thanksgiving Day football game is a great signal for the stock market. A great story? *I guess.* A great story, also backed by robust empirical evidence? *Not exactly.* And even if it were backed by evidence, you would be hard-pressed to raise investment capital to invest in the strategy.

But perhaps the "Dallas Cowboys" indicator is a bit far-fetched. How about the "52-week low" stock screen? Many of my stock-picking friends love this screen, thinking that 52-week low stocks are "cheap," on average, and therefore must offer the potential for great return relative to other stocks in the investment universe. Unfortunately, "52-week low stocks" are synonymous with what academic researchers call "low momentum stocks." Low momentum stocks, for those who shy away from reading academic finance journals, have been shown to be one of the worst performing groups[31] of stocks one can own.[32]

There are many other stock market superstitions—sell in May and go away; let your winners run, but cut your losses; head and shoulders patterns; this is a stock-pickers' market; invest in what you know; buy with a margin of safety; and so forth. Some of these stories are backed by evidence, others are not. The main point is that one's investment process should not be based on a story, but rather, on an evidence-based process that demonstrates robustness over time. Below, I outline 3 common stories (there are many more), where empirical evidence is lacking.

Story #1: Warren Buffett Beats Ben Graham

Ben Graham, Warren Buffett's mentor and original employer, had a strict focus on margin of safety. Graham's investment philosophy was to <u>always buy cheap</u> and never stray from a low price strategy. The essence of Ben Graham is captured in two of his recommended investment approaches:

1) Purchase stocks at less than their net current asset value, a strategy Graham considered "almost unfailingly dependable and satisfactory."[33]
2) Create a portfolio of stocks a minimum of 30 stocks meeting specific price-to-earnings criteria (below 10) and specific debt-to-equity criteria (below 50 percent).[34]

Both of these investment approaches maintain an overarching theme involving paying a low-price, independent of quality.

When Buffett came in to the spotlight, he suggested a wrinkle in Graham's original approach. Buffet's own words capture the flavor of his investment approach:

> *"It's far better to buy a wonderful company at a fair price than a fair company at a wonderful price."*[35]

In a Buffett world, Coke at a price-to-earnings ratio of 20 might be a value stock, but the textile firm Berkshire Hathaway may be overpriced at a P/E of 5. In a Graham world, Berkshire Hathaway is always the better bet. Anecdotally, it is easy to claim that Buffett was the clear winner in the horse race against Graham. But are we suffering from availability bias or story-bias when we make this conjecture?

What does the actual evidence have to say on the subject?

I can empirically verify whether a Buffett or Graham philosophy has been more effective over the past 37 years. To do so, I need to quantify Warren Buffett and Ben Graham's strategies in a systematic way. Joel Greenblatt, famous for his book, *The Little Book that Beats the Market,* tells a story about a systematic investment approach that encapsulates the Warren Buffet mantra of trying to "buy a wonderful business at a

---

[31] Jegadeesh, N., and S. Titman, 1993, "Returns to buying winners and selling losers: Implications for stock market efficiency," *Journal of Finance* 48, p. 65-91.

[32] This is not to say that 52-week low is ineffective in every situation, but one needs to be thoughtful when selecting a screening technology.

[33] Graham, B., and D. Dodd, *Security Analysis,* New York: McGraw-Hill, 1934.

[34] Graham, B., "A Conversation with Benjamin Graham," *Financial Analysts Journal* 32, p. 20–23.

[35] Buffett, W., "Chairman's Letter," Berkshire Hathaway Inc. Annual Report, 1989.

fair price." Greenblatt's formula is straightforward: Rank all stocks on their earnings before interest and taxes relative to their total enterprise value (EBIT/TEV). EBIT/TEV serves as the "cheapness" indicator for a given security (labeled "Graham" in Table 1). Next, measure the "quality" of a firm by calculating the ratio of EBIT to capital (labeled "Quality" in Table 1), which satisfies Buffett's own criteria that the "more appropriate measure of managerial economic performance is return on equity capital."[36] To generate the Warren Buffett clone strategy, we simply average the EBIT/TEV and EBIT/CAPITAL ranks and then purchases the top-ranked stocks based on the *combined* "cheapness" and "quality" ranking (labeled "Buffett" in Table 1).

Table 1 figures are from our book, *Quantitative Value.*[37] The performance metrics are calculated over the 1974 to 2011 timeframe and the universe consists only of investable firms (we eliminate small/micro caps). The far left column is the performance of Warren Buffett stocks as captured by Greenblatt's combined cheapness and quality measure. The second column represents the Graham cheap-stock strategy using only EBIT/TEV as the sorting variable. The third column is the stand-alone quality measure. The fourth column is the S&P 500 total return index. Each active strategy ranks stocks on the respective metric every June 30th and rebalances annually. The results reported represent the performance of the top decile of stocks for a given measure.

TABLE 1

	Buffett	Graham	Quality	S&P 500
CAGR	13.94%	15.95%	10.37%	10.46%
Std. Dev.	16.93%	17.28%	17.04%	15.84%
Downside Dev.	12.02%	11.88%	11.35%	11.16%
Sharpe	.55	.64	.35	.37
Sortino	.80	.96	.56	.56
Max Drawdown	-36.85%	-37.25%	-47.15%	-50.21%

The performance for the Buffett formula is admirable over the time period analyzed. Annual growth rates are almost 3.5% higher per year than the S&P 500 benchmark, and the Sharpe and Sortino risk-reward calculations are also stronger. But the Graham strategy outperforms on nearly every metric. The Graham strategy beats the market by over 5% a year, on average, and risk-reward metrics are much stronger than both the benchmark and the Buffett strategy. The evidence supports the argument that the original Graham value-investment philosophy is superior to the updated Buffet value-investment philosophy.

How is it possible that Graham beats Buffett? The answer lies in the quality component of the Buffett philosophy. If we examine the quality strategy's stand-alone performance we notice that the results are slightly weaker than the benchmark, suggesting that any strategy that moves out of cheap stocks with a quality component will correspondingly dilute overall performance. We see this borne out in the Buffett results, which represent a mix of a quality component and a low-price component. As we summarize in *Quantitative Value*, "[an equally weighted combination of quality and price algorithm] systematically overpays for quality. It is structurally flawed, leading us to fish in the wrong pond." The lesson from the evidence is that Graham was correct, on average. And yet, the story of value-investing has slowly evolved away from strictly buying cheap stocks to buying stocks across the price spectrum based on quality attributes that are not useful if a stock is not cheap. Unfortunately, this revamped value investing story is not backed by robust empirical evidence. Warren Buffett is merely an anecdote associated with a great story, but the tale told by Graham should maintain its status as the "golden rule of value investing."

Story #2: Economic Growth Drives Stock Returns

Should investors favor strong economic growth? Of course they should if they want to earn high returns. Strong growth drives profits, which drives returns.

FIGURE 17

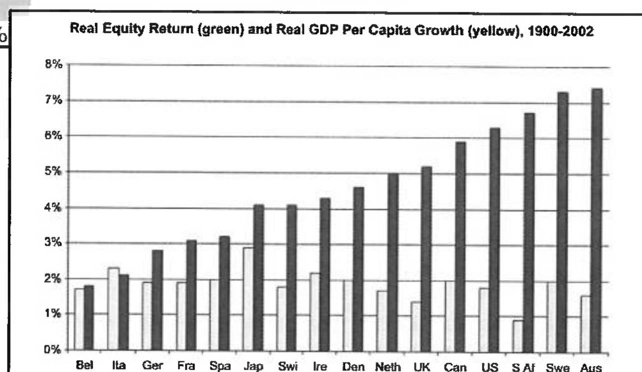


Fig. 1. Real equity return (green) and real GDP per capita growth (yellow), 1900–2002. Real return data (dividends and capital gains) are from Dimson et al. (2003). Real GDP per capita growth rate data are from Dimson et al. (2002), updated for 2001–2002 using data from *World Development Indicators*. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

If economic vitality didn't matter, all the time spent pontificating over economic figures and developing

[36] Buffett, W., "Chairman's Letter," Berkshire Hathaway Inc. Annual Report, 1977.

[37] Gray, W., and T. Carlisle, *Quantitative Value*: New York, John Wiley & Sons, 2012.

growth forecasts associated with these estimates would be a complete waste of time, right? *Not so fast*. I'm going to let you in on a secret: investors *focused on economic growth are wasting their time.*

Jay Ritter tells a compelling evidence-based story that economic growth *doesn't benefit* stockholders. [38] If anything, the evidence suggests a negative correlation between equity returns and GDP growth. [39] Figure 17 shows the relation between real equity returns and real per capita GDP growth for 16 countries over the 1900-2002 period—over a 100 year testing period!

The figure highlights the fact there is no relationship between stock returns and GDP growth. And yet, investors are so focused on the powerful narrative that GDP growth increases corporate profits, they forget to review the underlying theory or evidence sustaining this bogus story. First, from a theory perspective, the only way a firm increases stockholder value is by investing firm capital in positive net present value projects. And it is unclear why strong economic growth will contribute to a firm's ability to identify more, or higher yielding, investment projects in a competitive economy. Buffett made this point painfully clear in his famous 1999 article in *Fortune* magazine. First, the oracle rattles off a handful of transformative high-growth industries that translated into terrible investments (e.g., airlines, automobiles, radios, and televisions). He then leaves us with a profound statement that lays out a logical case that investors shouldn't fall in love with growth for growth's sake:

> *"The key to investing is not assessing how much an industry is going to affect society, or how much it will grow, but rather determining the competitive advantage of any given company and, above all, the durability of that advantage. The products or services that have wide, sustainable moats around them are the ones that deliver rewards to investors."* [40]

Buffet reminds investors why they shouldn't cling to macroeconomic growth stories. So, in which area should investors focus? As Ritter says quite succinctly: "current earnings yields." Translated for non-finance geeks, this simply means price. And as any intelligent investor will tell you, the price you pay has everything to do with the returns one will receive. If an investor pays a high price for a given asset, they can expect low returns; if the same investor pays a low price for a given asset, they can expect high returns. The real story here is that high equity returns are earned by investors who focus on paying low prices for firms with strong abilities to invest in positive net present value projects. It may be that the best prices can be had in times of low economic growth, whereas we tend to overpay in a growing economy. The idea that strong economic growth translates into strong stock returns is a superstition, not backed by evidence.

Story #3: The Payout Superstition

Every quarter, boards across America wrestle with the complex question of dividend policy. Perhaps the company has excess cash that should be paid out as a dividend? Or perhaps cash should be directed to high net-present-value projects? It's a nuanced and sophisticated debate, which makes it the perfect breeding ground for generating investor superstitions.

Quant heavyweights Cliff Asness (AQR) and Rob Arnott (Research Affiliates) have noted that market observers often predict that low dividend payout ratios imply higher earnings growth in the future. [41] Conversely, when dividend payout ratios are high, commentators suggest that earnings growth will slow in subsequent years. I call this story the "payout superstition."

Again, the logic seems to make sense: if companies retain earnings (i.e., low dividend payout) and plow them back into promising projects, earnings growth should be higher in the future; conversely, if companies don't see any growth opportunities, they will push cash back to shareholders (i.e., high dividend payout) and future earnings shouldn't experience robust growth.

---

[38] J.R. Ritter, 2005, "Economic Growth and Equity Returns," *Pacific-Basin Finance Journal* 13, p. 489-503.

[39] Ritter find a cross-sectional correlation of -.37 for the compounded real return on equities and the compounded growth rate of real per capital GDP for 16 countries over the 1900-2002 period.

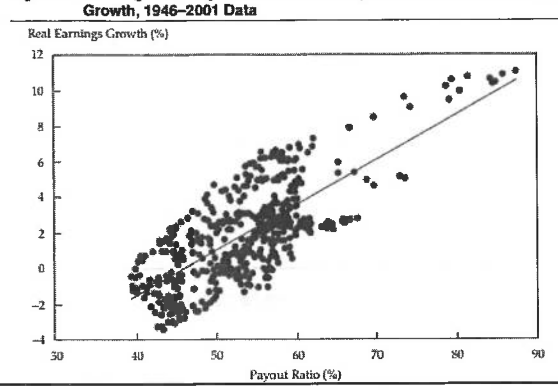
[40] Mr. Buffett on the Stock Market, Warren Buffett and Carol Loomis. http://money.cnn.com/magazines/fortune/fortune_archive/1999/11/22/269071/, accessed 2/10/2014.

[41] Arnott, R., and C. Asness, 2003, "Surprise: Higher Payout Rates = Higher Growth Rates," *Financial Analysts Journal* 59, p. 70-87.

ARE YOU TRYING TOO HARD?

FIGURE 18

Figure 2. Scattergram of Payout Ratio vs. Subsequent 10-Year Real Earnings Growth, 1946–2001 Data



The payout superstition is a great story, but is this really how the world works from an empirical standpoint?

Arnott and Asness looked at historical payout ratios and earnings growth of stocks broadly representative of the market. Figure 18 is a scatterplot showing payout ratios and subsequent ten-year real earnings growth from 1946-2001.

The evidence indicates there is a *positive* relationship between payout ratio and future earnings growth. That is, higher dividend payout ratios predict higher growth, while lower payout ratios predict lower growth—the opposite of the payout superstition. As my 2-year old says after dropping food from the table: *"Uh-oh."*

Asness and Arnott suggest a few hypotheses for why this might be true:

- Since managers don't like to cut dividends, if they are concerned about the sustainability of earnings in the future they would not offer a higher dividend today; a higher payout ratio is a signal that they think future prospects look poor.
- When earnings are not paid out, cash is used to finance poor investments (malinvestment), leading to reduced earnings growth.
- When managers hold cash, it may signal "empire building," where managers try to increase their power, rather than act to benefit shareholders.

Arnott and Assness's analysis suggests that if a firm has extra cash, there are reasonable arguments why they should pay out cash as dividends, rather than hold it or invest it in disastrous projects that could destroy value. In a world where malinvestment and empire building are pervasive, dividends might provide a valuable signal about a firm's shareholder policies. And perhaps mischievous corporate managers are exploiting the payout superstition for their own benefit?

The Moral of the Stories

The number of fairy tales, rules of thumb, and other sorcery sold in the financial markets are too numerous to list. I've highlighted 3 of the more coherent and believable stories in the marketplace that are called into question by empirical footings. The lesson is clear for all of us who enjoy a great investment pitch: In order to be good investors, we need to appreciate our natural preference for coherent stories, and our innate dislike for evidence. Don't be the pigeon doing a "pellet voodoo dance," when it has already been shown that the pellet voodoo dance doesn't work.

SECTION 5C: OVERCONFIDENCE

> *"We are prone to overestimate how much we understand about the world and to underestimate the role of chance in events."*
>
> *--Dan Kahneman, Thinking, Fast and Slow.*

Overconfidence, or the inability to appropriately calibrate our forecasts, is often cited as among the most robust empirical finding in psychology experiments.

Let's try a game.

Spend a couple of minutes identifying a low and high value answers to the questions in Figure 19, such that you are 90% confident the answer lies in between your upper and lower bound. To be clear, answering "negative infinity" and "positive infinity," while clever, is missing the point of the game. You want to calibrate your upper and lower bound appropriately: not too cold, not too hot—just right. Go for it.

FIGURE 19

Question	Low	High
1. Martin Luther King's age at death		
2. Length of the Nile River		
3. Number of countries that are members of OPEC		
4. Number of books in the OldTestament		
5. Diameter of the moon		
6. Weight of an empty Boeing 747		
7. Year in which Wolfgang Amadeus Mozart was born		
8. Gestation period (in days) of an Asian elephant		
9. Air distance from London to Tokyo.		
10. Deepest (known) point in the oceans.		

If you are like most people who play this game, you are reliably overconfident. My own ad-hoc experimental evidence (around 1,000 participants) on this questionnaire is that individuals typically get 3/10 correct, when a well-calibrated individual should on average get 9/10 correct, consistent with the confidence interval. This low-scoring result holds even when I warn test-takers that prior test-takers have been systematically overconfident in their upper and lower bounds. I then encourage them to increase the bounds of their ranges. Of course, nobody listens, and on average, only 3/10 of the correct answers actually sit within the individual's confidence range. We are hard-wired to be overconfident. [42]

What cognitive biases are causing us to be overconfident? One answer may lie in our human desire to pursue and misuse useless information. Our brains immediately interpret more information as better information, which leads to more confidence, with no corresponding increase in forecast accuracy.

There is a great study by Bastardi and Shafir appropriately titled, "On the pursuit and misuse of useless information." [43] The paper is filled with experiments that show our brain's inability to properly process information in a variety of circumstances. The abstract of the paper says it best:

> *"Decision makers often pursue noninstrumental information--information that appears relevant but, if simply available, would have no impact on choice. Once they pursue such information, people then use it to make their decision. Consequently, the pursuit of information that would have had no impact on choice leads people to make choices they would not otherwise have made."*

Here is an example experiment from Bastardi and Shafir's research. The authors ask different groups to make a decision on a mortgage application. One group is faced with the following information set:

Group 1: Imagine that you are a loan officer at a bank reviewing the mortgage application of a recent college graduate with a stable, well-paying job and a solid credit history. The applicant seems qualified, but during the routine credit check you discover that for the last three months the applicant has not paid a $5,000 debt to his charge card account.

- ***Do you approve or reject the mortgage application?***

Group 1 approves only 29% of the applications and rejects 71%.

The authors play a trick on the second group by leading them to believe they have more information. The hypothesis is that the subjects will interpret their supposed "special information" as information that can lead to a more accurate decision.

Group 2: Imagine that you are a loan officer at a bank reviewing the mortgage application of a recent college graduate with a stable, well-paying job and a solid credit history. The applicant seems qualified, but during the routine credit check you discover that for the last three months the applicant has not paid a debt to his charge card account. The existence of two conflicting reports makes it unclear whether the outstanding debt is for $5,000 or $25,000, and you cannot contact the credit agency until tomorrow to find out which is the correct amount.

- ***Do you approve or reject the mortgage application or wait?***

Only 2% of the respondents approve the application, while 23% reject the application, and a majority (75%) chooses to wait for the additional information. For the majority who wait a day to get the additional information, the authors present them with the following tidbit:

Next day they find out the amount is $5,000

- ***Do you approve or reject the mortgage application?***

For the majority who held out, 72% approve the application and 28% reject the application. In sum, for group 2, 54% approve the application and 21% reject the application. The approve rates are substantially higher than for group 1.

---

[42] The answers are below:

1. 39 yrs
2. 4,187 miles
3. 13 countries
4. 39 books
5. 2160 miles
6. 390,000 pounds
7. 1756
8. 645 days
9. 5,959 miles
10. 36,198 feet

http://www.tim-richardson.net/misc/estimation_quiz.html. Accessed, 2/10/2014.

[43] C. Bastardi A, and E Shafir, 1998, "On the Pursuit and Misuse of Useless Information," *Journal of Perspectives on Social Psychology* 75, p. 19-32.

ARE YOU TRYING TOO HARD?

What is going on in this experiment? The authors have effectively given group 1 and group 2 the *exact same information set,* but because it is meted out over time, the second group perceives they have more information, which changes their decision-making process. Humans are cognitively inclined to overvalue information that requires effort or time to obtain.

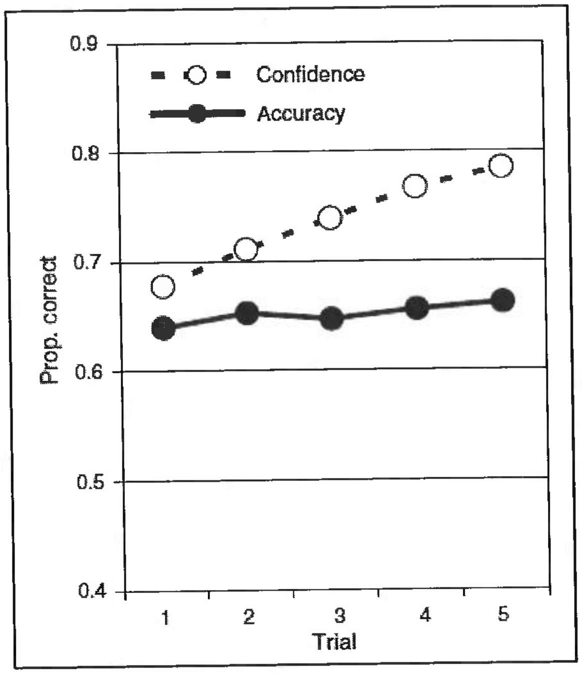
To make the point more vividly that more information doesn't necessarily translate into better-decision making, Professor Claire Tsai and colleagues directly test the relationship between information and forecast accuracy.[44] The "rational" hypothesis suggests that each information piece received will be appropriately weighted and integrated into a forecast. Confidence in the forecast will be updated via appropriate statistical means (i.e., Bayesian updating). The "irrational" hypothesis is that humans will interpret more information as better information, without considering whether the additional information actually enhances their forecast ability. The prediction is that forecast accuracy will not improve as people receive more information, but their confidence in their forecast will increase linearly with more information.

The authors collect subjects who self-identify as being knowledgeable about college football. They present different subjects with up to 30 data points. The subjects are then asked to present a forecast for football game outcomes and their confidence in their forecast. To spice things up a bit, the researchers give the subjects the information set in such a way that they receive the most predictive pieces of information first and each subsequent piece of information is less and less useful for predicting football game outcomes.

The results of the study support the human hypothesis and reject the econ hypothesis. Humans incorrectly interpret more information as better information. Their forecast accuracy does not improve with more information, but their confidence in their forecast grows linearly with the amount of information received.

[44] Claire Tsai, Josh Klayman, and Reid Hastie, 2008, 'Effects of amount of information on judgement accuracy and confidence," *Organizational Behavioral and Human Decision Processes* 107, p. 97-105.

FIGURE 19

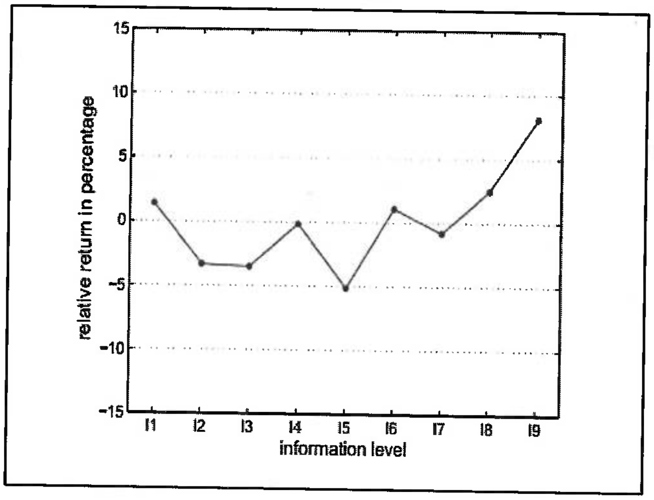


A team of theoretical physicists and social scientists examine the effect of information on forecast accuracy and confidence.[45] The researchers create an experimental trading lab where traders are randomly given different information sets with nine different levels of information. I1 has no information; I9 is essentially an insider, with a nearly perfect information set. Once traders receive their information sets, they play a live trading game where the subjects try to maximize their returns. The intent of the experimental design is to capture an element of the real-world marketplace where some traders are better informed than others and these traders trade with each other in financial markets. Amazingly, only the most informed traders with complete insider information can reliably beat the market. While it is unsurprising that a total insider could beat the market, it is striking that the partially informed traders do not. While the partially informed traders should outperform the market because they have privileged information, they do not largely because they suffer from overconfidence and overvalue their own information set, and therefore can't use it effectively. In fact, partially informed traders on average, underperform. Uninformed investors, who know they have no information, are less likely to suffer from the

[45] Bence T., Scalas E., Huber J., and M. Kirchler, 2007, "The value of information in a multi-agent market model," *European Physical Journal B* 55, p. 115-120

cognitive bias of overconfidence, and thus end up achieving the market return, which in this experimental setting, is an admirable achievement.

FIGURE 20



What are we to make of this study? Is the takeaway that insider information is helpful for generating market-beating returns? Well, *no kidding.* In a game setting, we would all trade on insider information if we had it. However, in the real-world, we must weigh the benefits of beating the market relative to the direct costs of being holed up in a prison cell wearing an orange jump suit and the indirect costs of having to shamefully stare into the mirror every morning.

And for the rest of the traders in the marketplace with partial information sets? These investors should be wary of interpreting their information as a way to enhance performance. If the experimental evidence is any guide, it is more likely that additional information is causing us to make *worse* decisions, not better decisions. We must always ask ourselves if the information we are collecting is adding value by enhancing our forecasts or detracting from value by making us more overconfident.

In addition to correctly assessing the value of information we receive, we must avoid self-attribution bias, which is our innate desire to attribute good outcomes to skill and bad outcomes to back luck. A good trade does not necessarily imply that the investor was better at "doing their homework," nor does a bad trade necessarily imply that the investor did a poor job of "doing their homework." And yet, our minds will often attribute the successful trade to our wonderful ability to collect and interpret mounds of filings submitted to the SEC, our ability to do the "hard work" of conducting channel checks on suppliers and customers, and our magnificent skill in being better than the average investor in the marketplace. Of course, when we endure a poor trade, we don't attribute the bad performance to a lack of skill, but instead, the bad trade can be squarely blamed on bad luck: an unforeseen change in government policy, or perhaps a remarkable change in the price of underlying commodities that "nobody could have seen coming."

A better approach for dealing with success and failures is to systematically discount success and overemphasize failures. Flip self-attribution bias on its head, or as Charlie Munger, Vice Chairman of Berkshire Hathaway, is often attributed as saying, "Invert; always invert." While unappealing to most, reiterating that we are not as smart as we thought we were and realizing the pain of bad decisions can actually make us stronger, since it is a more accurate representation of reality. As we say in the Marine Corps, "Pain is weakness leaving the body."

The potent combination of overvaluing additional information and self-attribution bias contribute to systematic and predictable overconfidence for discretionary decision-makers. This overconfidence leads to value-destroying decisions in the context of financial markets.

SUMMARIZING WHY EXPERTS FAIL TO BEAT MODELS

The expert hypothesis is based on the following flawed assumptions:

- Qualitative information increases forecast accuracy.
- More information increases forecast accuracy.
- Experience and intuition enhance forecast accuracy.

The assumptions underlying the expert's hypothesis are empirically invalid because: "Soft," or qualitative, information doesn't enhance forecasting ability; more information doesn't enhance forecasting ability; and experience doesn't enhance forecasting ability.

Systematic models work because the human mind is reliably unreliable.

Let us push reality aside for a moment, and make the claim that most of us are truly evidence-based decision-makers who are not influenced by stories that capture our imagination and impact our decision-making ability. If we are truly empirical-based individuals, the evidence overwhelmingly suggests that we should all be using models and other algorithms to make decisions, rather than relying on experts.

ARE YOU TRYING TOO HARD?

But who is ready to concede that a machine is better at making decisions? You are probably like me: the idea of scrapping my years of hard-won experience is awfully hard to swallow. (see Figure 20.[46])

FIGURE 20

Humans naturally seek to fulfill what Maslow—famous for developing the human hierarchy of needs[47]—calls our innate need for esteem and self-actualization. We want to feel that our opinions and judgment matter. Recognizing the fact that simple models outperform experts directly challenges our self-directed desire to achieve goals, gain confidence, and feel a sense of achievement. We want to feel like our efforts are worthwhile, but we often devote little effort to understanding if our frenetic activity actually adds value.

Consider the act of banging one's head against the wall for 10 hours a day, 7 days a week. Banging your head against the wall involves a lot of activity, but because the outcome of this activity is clearly "bad," it is easy to know that this focused effort is a waste of time. However, what if we are spending 10

hours a day contacting CEOs about the prospects of their companies? Is this intense activity valuable? Are we learning anything that is actually helping us? A lot of investors assume it is, but have they ever systematically reviewed this assumption? Maybe this, or other, so-called "value-add" activities performed by experts is equivalent to banging one's head against the wall? Perhaps the activities are detracting from value, not contributing to value?

I can't say with certainty, but based on the bevy of tests cited above, I can conjecture that while the analyst is clearly collecting more information, the information may do nothing to enhance the analyst's forecasting ability. In fact, it is quite likely that the additional information *detracts* from our ability, as the analyst becomes systematically overconfident in his forecast of the future. Overall, any potential information edge that may exist is overwhelmed by costs associated with cognitive bias issues.

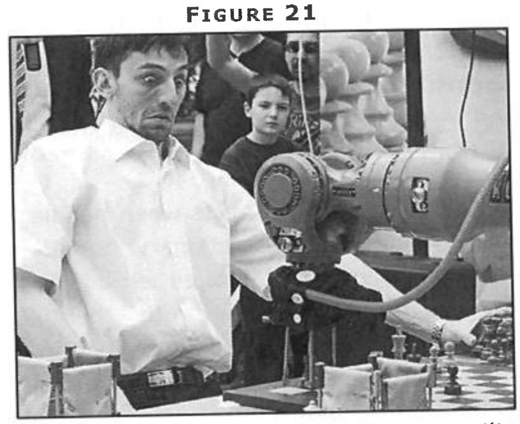
[16] Image source unknown.

[47] Maslow, A.H. (1943), "A theory of human motivation," *Psychological Review* 50, p. 370–96.

SECTION 6: WHY NOT USE MODELS?

Image you are watching Gary Kasparov, Russia's preeminent chess master, taking on IBM's Deep Blue, a cold, calculating box designed by a bunch of geeks. (See Figure 21[48]).

FIGURE 21



During the match, Gary is sweating it out, smiling when he makes a nice move, and cringing in pain when Deep Blue takes his queen. We see that Gary is like us. He is *familiar*; the machine is just an inhuman metal box. The machine has no emotion, no feelings, no empathy. Who do we want to win the match? We want Gary. He's like us and we have a preference for the familiar (yes, another bias, I know). Nobody wants a computer to win.

And so what if the machine is actually better at chess than a human? We get it: Deep Blue with its ability to analyze 200 million positions per second, can best a human opponent. Does that mean we want a chess-playing computer mainframe making our life and death medical decisions—*even if the evidence suggests we should?* Humans might be willing to put up with a flawed, but familiar human, because we empathize with flesh and blood. If the heart surgeon kills my aunt because he accidently tied the tubes the wrong way, that's unfortunate, and I'm angry, but "people make mistakes, we're all human." But imagine if a robot performs surgery on my aunt and she dies because the robot tied the tubes the wrong way. My immediate

[48] Image source unknown.

reaction: "Who in the heck thought it was a good idea to have a robot perform heart surgery—Where's my lawyer!" However, the truth is, the robot is much less likely to make such a mistake, on average. We should be rooting for machines that make fewer errors, not excusing human error.

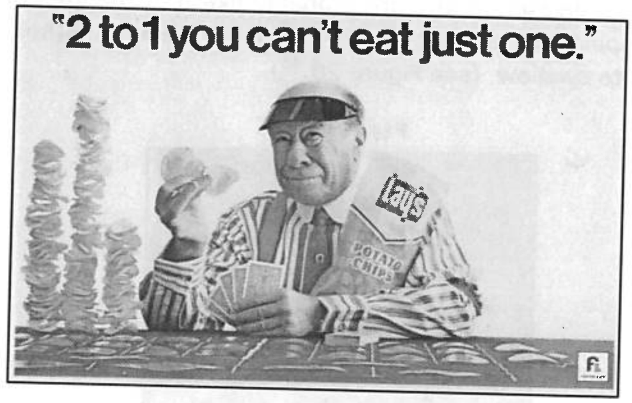
Even if one buys the argument that models can be useful, one might object that models are too limited and cannot be applied in sophisticated contexts like investment decision making. What, for example, is the algorithm going to say when we face a unique situation the world has never seen? This time is *actually* different.[49] The story is that the human expert can adapt and create on-the-fly modifications to the model that create value. This well-trodden, but empirically busted, rebuttal against algorithms is deemed the "broken leg theory" and relies on the false premise that humans don't suffer from System 1 flaws.

Consider an algorithm that predicts when people will go to the movie theater. A discretionary human expert identifies that someone has a broken leg and is able to update the quantitative model with their "special knowledge" and beat the model. Sounds good in theory; however, as the experimental psychology research shows, humans are unable to properly differentiate between information that actually enhances forecast accuracy and the reams of additional useless information they believe can enhance forecast accuracy. What about the upcoming weather patterns suggesting that rain is imminent? Or what about the fact that the star in the movie was just charged with a DUI on the nightly news? This must matter for predicting movie attendance—or does it?

Without the evidence, and only a story to rely on, we are on shaky ground. In summary, discretionary decision makers are often able to identify the value-enhancing modifications that can theoretically outperform a simple model, however, they simultaneously identify value-destroying modifications that cause them to underperform. Discretionary experts' inclination to "modify" simple models resembles a bag of Lay's Potato Chips—*the experts can't eat just one modification.*

[49] E.g., "The fed has never intervened in the market like this."

FIGURE 22



Still think that you are the exception who can reliably add value? **<u>We ALL believe we are better than average</u>**. The crushing reality is "You are less beautiful than you think."[50]

- Are you a better driver than average? *93% of US citizens think so too.*[51]
- Are you a great teacher? *94% of professors think so too.*[52]
- Are you a better than average stock-picker? *Of course you are.*

Relegating your decision-making processes to systems requires a massive dose of humble pie. (see Figure 23).

Most—if not all—are unable to consume this dish. But to be a better decision maker we must eat our humble pie. As I have shown in this essay, *<u>in order for decision making to be effective, it must be systematic.</u>* And the only systematic thing about humans is our flaws. Therefore, it is best to leave the stock picking to Warren Buffett, and for the rest of us, who suffer from behavioral biases, which result in flawed decision making, we should stare into the mirror, and ask ourselves: ***Are You Trying Too Hard?***

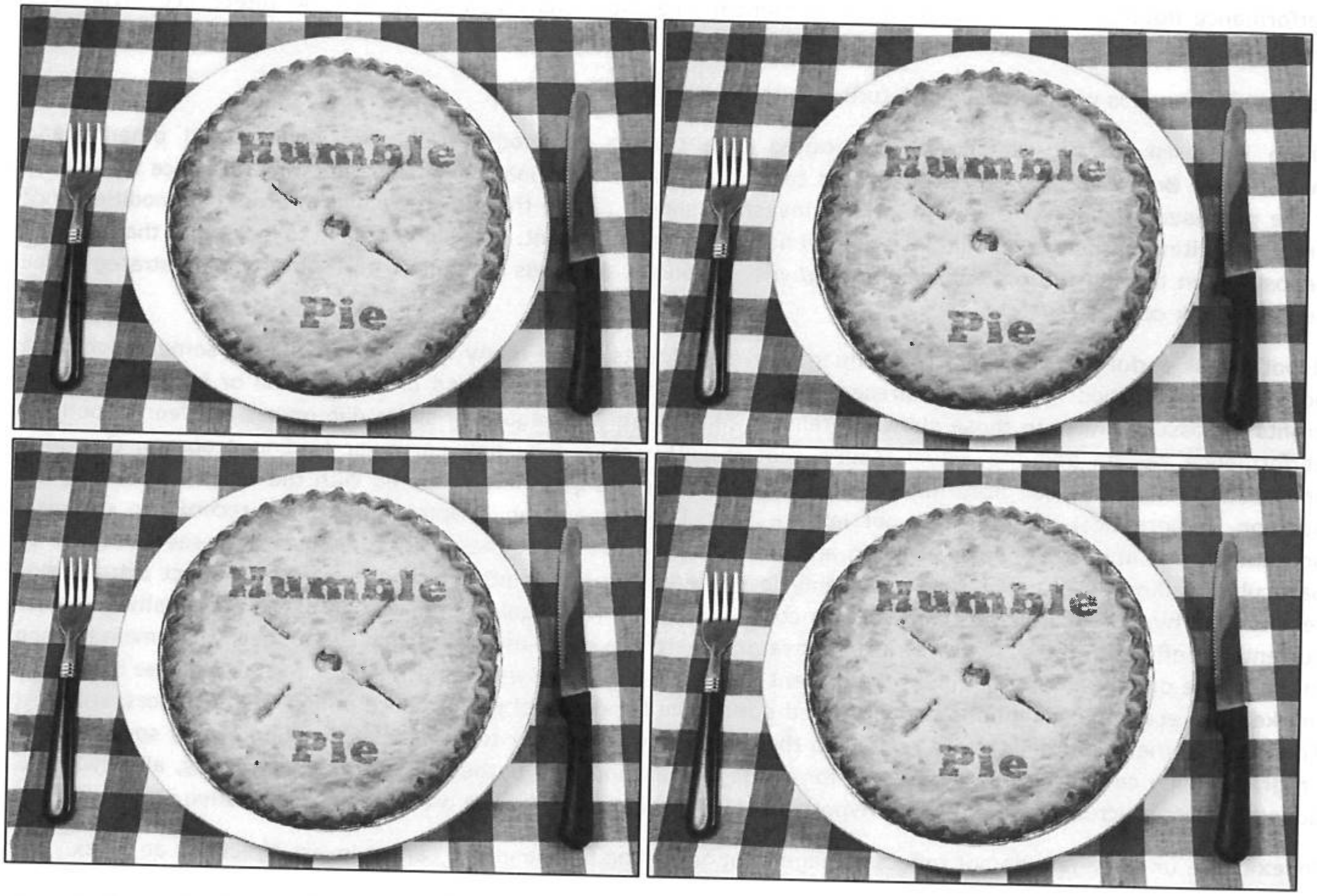
[50] http://www.scientificamerican.com/article/you-are-less-beautiful-than-you-think/, accessed 2/10/2014.

[51] Svenson, O., 1980, "Are we all less risky and more skillful than our fellow drivers?" *Acta Psychologica* 47, p. 143-148.

[52] Cross, P., 2006, "Not can, but will college teaching be improved?" *New Directions for Higher Education* 17, p. 1-15.

Are You Trying Too Hard?

FIGURE 23



DISCLOSURES

Performance figures contained herein are hypothetical, unaudited and prepared by Alpha Architect, LLC; hypothetical results are intended for illustrative purposes only.

Past performance is not indicative of future results, which may vary.

There is a risk of substantial loss associated with trading commodities, futures, options and other financial instruments. Before trading, investors should carefully consider their financial position and risk tolerance to determine if the proposed trading style is appropriate. Investors should realize that when trading futures, commodities and/or granting/writing options one could lose the full balance of their account. It is also possible to lose more than the initial deposit when trading futures and/or granting/writing options. All funds committed to such a trading strategy should be purely risk capital.

Hypothetical performance results (e.g., quantitative backtests) have many inherent limitations, some of which, but not all, are described herein. No representation is being made that any fund or account will or is likely to achieve profits or losses similar to those shown herein. In fact, there are frequently sharp differences between hypothetical performance results and the actual results subsequently realized by any particular trading program. One of the limitations of hypothetical performance results is that they are generally prepared with the benefit of hindsight. In addition, hypothetical trading does not involve financial risk, and no hypothetical trading record can completely account for the impact of financial risk in actual trading. For example, the ability to withstand losses or adhere to a particular trading program in spite of trading losses are material points which can adversely affect actual trading results. The hypothetical performance results contained herein represent the application of the quantitative models as currently in effect on the date first written above and there can be no assurance that the models will remain the same in the future or that an application of the current models in the future will produce similar results because the relevant market and economic conditions that prevailed during the hypothetical performance period will not necessarily recur. There are numerous other factors related to the markets in general or to the implementation of any specific trading program which cannot be fully accounted for in the preparation of hypothetical performance results, all of which can adversely affect actual trading results. Hypothetical performance results are presented for illustrative purposes only.

Indexes are unmanaged, do not reflect management or trading fees, and one cannot invest directly in an index.

There is no guarantee, express or implied, that long-term return and/or volatility targets will be achieved. Realized returns and/or volatility may come in higher or lower than expected.