The Effect of Index Fund Competition on Money Management Fees

Yang Sun[†]

Abstract

This paper estimates the effect of competition from low-cost index funds on fees in the money management industry. A difference-in-differences analysis exploiting the staggered entry of index funds finds that while actively managed funds sold directly to retail investors reduce fees by six percent, those sold through brokers *increase* fees by four percent. Additionally, actively managed funds, especially closet indexers, shift away from holding the index portfolio. A market segmentation model illustrates that beyond a price-competition effect, the index fund entry creates a selection effect that isolates the least-price-sensitive investors and results in a price increase for this group.

[†]MIT Sloan School of Management (email sunyang@mit.edu). I am indebted to Nittai Bergman, Andrey Malenko, Stewart Myers and especially Antoinette Schoar for their invaluable guidance and encouragement. I thank Jean-Noel Barrot, Asaf Bernstein, Tara Bhandari, Hui Chen, Martijn Cremers, Pengjie Gao, Xavier Giroud, Xing Hu, Raj Iyer, Dirk Jenter, Li Jin, Mark Kritzman, David Laibson, Chen Lin, Tse-Chun Lin, Hong Luo, Marco Di Maggio, David Matsa, William Mullins, Jun Pan, Jonathan Parker, Tarun Ramadorai, Jonathan Reuter, Nancy Rose, Stephen Ross, Felipe Severino, Pian Shu, Dragon Tang, Adrien Verdelhan, Jiang Wang, Xiao Yu Wang, Keith Wong, Yan Xu, as well as seminar participants at MIT Finance Lunch, MIT Finance Seminar, University of Hong Kong and University of Notre Dame for helpful comments. In addition, I am grateful to Gary Gastineau, John Minahan, Kim Perry, Gita Rao, Steve Snider and Brian Walters for sharing insights on the institutional setup. All errors are mine.

1 Introduction

The fees charged by money managers, particularly in the retail sector, have been a subject of academic debate for at least half a century (Wharton (1962)).[1] The topic has recently received renewed attention in several studies that have documented the amount of fee revenues earned in the finance industry (French (2008); Philippon (2012); Greenwood and Scharfstein (2013)). The authors conjecture that the gradual decrease in the average mutual fund fees probably reflects the competitive pressure from the entry of new and cheaper investment vehicles, such as index funds and exchange-traded funds (ETFs) (see, also, Cochrane (2013)). However, these studies suggest that the fee decline is surprisingly slow, especially considering the intensity of the competition.

The finance literature has pointed to several potential explanations for the slow decline in relatively high mutual fund fees. One argument can be drawn from the theory that fees are a reflection of managerial abilities (Berk and Green (2004); Berk and Van Binsbergen (2012)): Persistent fees suggest that talent, as a fraction of assets under management, has not declined. An alternative argument is that the increase in supply coincides with increased investor demand (Cochrane (2013)). While these explanations are plausible, it remains an open question how the price setting mechanism in the money management industry responds to the competitive pressure from passive funds.

This paper focuses on identifying the effects of passive fund entry on fees and investment strategies in the mutual fund industry. Exploiting the staggered entry of low-cost index funds and ETFs as a novel natural experiment, the analysis shows that the competition from passive funds affects average mutual fund fees in the market only slightly; but, importantly, there are strong divergent effects across market segments. In response to the entry

[1] See Coates and Hubbard (2007) for a full review.

of index funds, active funds sold directly to investors (direct channel funds) reduce fees by six percent, consistent with the standard price competition effect. However, the funds sold by intermediaries and bundled with financial advice (broker channel funds) *increase* fees by about four percent. This result runs counter to a basic model of competition in which the new entrant uniformly exerts downward price pressure on all incumbents. Because the effects in the two channels go in opposite directions, the time series of aggregate fees demonstrates only small impact from the entry. On average, actively managed funds lowered fees by only about 0.2 percent in response to index fund entry. Hence, taking account the heterogeneity across the fund distribution channels may help explain the empirical puzzle that average fees decline slowly in spite of sharply increasing competition.

The empirical analysis exploits the time series variation in the entry of low-cost index funds and ETFs into different U.S. equity fund investment styles. In particular, the study uses the sequence of Vanguard entry as a proxy for competitive shocks from index funds. During the sample period 1970-2005, Vanguard entered eight styles[2] at a fraction of the fees compared with incumbent actively managed funds. Taking advantage of the staggered entry, the regression follows a difference-in-differences (DID) specification, and it allows differential fee responses across distribution channels. In this panel DID, causal inference requires that the specific timing of the entry is uncorrelated with the outcome variables – this assumption is supported by anecdotal and econometric evidence.

The DID results show that the entry of index funds leads to significant divergent reactions across retail fund distribution channels. Active fund share classes[3] in the direct channel reduce fees by about seven to nine basis points, or six to seven percent of their mean level. In contrast, share classes in the broker channel increase fees by nine to ten basis points, or

[2]Styles are measured using Lipper portfolio-based classification.

[3]A share class is a type of mutual fund share with a distinct fee structure.

four to five percent of the mean. An examination of the dynamics suggests that the changes take about three years to stabilize and are long-lived. The results are robust to allowing the observable fund characteristics to affect the funds' responses to competition. Further, the results are supported by a list of robustness checks and falsification tests.

To explain the findings, especially the divergent price responses across the distribution channels, I build a model that shows how mutual fund companies optimally use the provision of financial advice as a screening device to price discriminate among investors, before and after the entry of lower-cost competitors. The model follows the setup in Ellison (2005) and hinges on the assumption that investors have heterogeneous price sensitivity. This assumption reflects the evidence in the literature that investors in the broker channel have little knowledge of the fees they pay (Capon et al. (1996); Alexander et al. (1998)). In the model, fee-sensitive investors self-select to the direct channel to purchase the fund alone, while fee-insensitive investors select to the broker channel for the fund-advice bundle. From this equilibrium, the entry of a low-cost index fund has two effects. First, there is a price-competition effect which presses average fees down. Second, there is a "*selection*" effect – i.e., the entry disproportionately draws demand from price-sensitive investors. Facing the new investor distribution which is skewed toward the less-price-sensitive types, actively managed funds optimally choose to retain the least-price-sensitive investors in the broker channel and charge them a higher fee.

The model generates additional predictions which are then tested in the data. First, consistent with the idea of asset attrition, the analysis shows that fund flows to the broker channel are negatively affected by the index fund entry – the average annual flow rate decreases by three to five percent. Second, the model predicts that the investors who purchase funds through the broker channel are less sensitive to fees. To test this notion, the sensitivity of fund flows to fees across the channels is estimated, controlling for past performance. The result shows that fund flows are significantly more sensitive to fees in the direct channel than in the broker channel. Third, the selection mechanism implies that the fee sensitivity in the broker channel should go *down* post the index fund entry, and an estimate on the change in fee sensitivity offers support for this hypothesis. Fourth, the model implies that actively managed funds engage in product differentiation to soften price competition from index funds. On this implication, the analysis finds that the level of active management increases, and the effect is concentrated in the funds that ex ante hold portfolios similar to the benchmark ("closet indexers").

The study further examines the staggered ETF entry into the passive fund market of U.S. equity funds. The ETF entry is a direct competitive threat for traditional index funds. In response to ETF entry, the share classes of index funds in the direct channel lower fees by 11 percent of the mean, and those in the broker channel *raise* fees by about 12 percent. The result closely resembles that from the Vanguard study and reinforces the main findings of this paper.

This paper contributes to the literature on the role of competition in the money management industry. Several empirical studies have examined the relationship between mutual fund competition and fees. Coates and Hubbard (2007) and Khorana and Servaes (2012) show that fees are negatively correlated with fund size and market share. Cremers et al. (2011) documents that in countries with explicit index funds, active funds are more active and charge lower fees. Wahal and Wang (2011) finds that in response to new fund entry, incumbents reduce management fee but increase distribution fee.

The present study makes two contributions to this literature. First, it provides a nuanced view on the effect of competition in the mutual fund industry, and this view offers a potential explanation for why aggregate mutual fund fees decline only slowly. By exploiting the heterogeneous competitive dynamics across distribution channels, the analysis shows that competition from low-cost financial innovations can hurt the high-price end of the market while benefiting the low-price end. The intuition highlighted here is related to Christoffersen and Musto (2002), who finds that money market funds increase fees after bad performance, because the remaining investors are the self-selected performance-insensitive types. Similar selection effects have been studied in other markets, for example, the home-mortgage market (Schwartz and Torous (1989)) and the pay-day loan market (Melzer and Morgan (2012)), though the literature has not documented evidence on divergent price responses to competition.

Second, the empirical strategy addresses an identification challenge that arises because cross-sectional correlations between competition and fees can be confounded by endogeneity or unobservable market characteristics. With evidence supporting the quasi-exogenous timing of the entry, the study argues that the estimated effects are *caused* by the competitive pressure from low-cost index funds and ETFs. The use of this natural experiment is a novel contribution to the literature, and it can be applied to study other effects of competition in the money management industry.[4] For example, Célérier and Vallée (2014) adopts the instrument I developed in this present paper and applies it to the context of European structured product market. The authors show that increased competition from the entry of ETFs leads to increased financial complexity. The focus of their paper is on product complexity, while the present analysis focuses on heterogeneous price responses across market segments.

This article also contributes to the study of market segmentation in the mutual fund industry. In this strand of the literature, Del Guercio and Reuter (2013), Del Guercio et al.

[4]Similar identification strategy has been used to study the causal effect of competition in the supermarket industry; see, for example, Matsa (2011).

(2010) and Christoffersen et al. (2013) study the different incentives for fund companies across the distribution channels. In addition, Bergstresser et al. (2009), Gennaioli et al. (2012), Mullainathan et al. (2012), and Chalmers and Reuter (2013) shed light on the demand for financial advice. Few studies have explicitly modeled the market segmentation as a price discrimination mechanism. The model in this paper shows that price discrimination can explain not only the segmentation of the mutual fund market by the sales channels, but also the divergent effects of passive fund competition.

Furthermore, the present work extends previous research on frictions in investors' behavior. The literature has examined the search costs (Sirri and Tufano (1998), Elton et al. (2004) and Hortacsu and Syverson (2004)) and bounded rationality (Barber et al. (2005), Gil-Bazo and Ruiz-Verdú (2008) and Choi et al. (2010)) of mutual fund investors. Recent theoretical development in behavioral economics also shows that companies can strategically increase complexity or adopt "shrouded" charges to extract information rents from consumers (Gabaix and Laibson (2006); Ellison (2005); Carlin (2009); Carlin and Manso (2011) and Gabaix et al. (2013)). The characteristic of fee-insensitivity highlighted in the present paper can be flexibly interpreted as a consequence of search cost, bounded rationality, or obfuscation. The analysis demonstrates that the frictions along these dimensions can have important equilibrium implications for the welfare effects of competition.

The remainder of the paper is organized as follows. Section 2 provides background information on the index funds and ETFs. Section 3 presents the theoretical framework and the model. Section 4 describes the data. Section 5 gives an overview of the empirical framework. Section 6 presents the results, and Section 7 concludes.

2 Background on Index Funds and ETFs

2.1 Index Funds

An index fund is a type of mutual fund whose investment objective is typically to achieve approximately the same return as a particular market index, either through directly investing in the securities included in the index or through replicating the index returns with synthetic methods. An index fund substantially reduces management costs due to relatively fixed portfolio holding, less trading and lower taxable realized capital gains. Vanguard launched the first broad-market index fund for retail investors in 1976 in response to the advocacy by academics,[5] and the company, especially the CEO John Bogle, has been a major force in promoting passive investment strategies. The market share growth of index funds became exponential in the mid-1980s (Figure 1, solid line). By 2011, the market share for broadly diversified open-end equity index funds had reached 22 percent. Table 1, Panel A lists the timing of the initial Vanguard index fund entry into various investment styles. In almost all investment styles, Vanguard was the first index fund to enter the market.[6] For this reason, the series of Vanguard entry is used to proxy for the timing of the competitive shocks from index funds.

[5]See, for example, Malkiel (1973) and Samuelson (1974). The first index portfolio was launched in 1971 by Wells Fargo for a single pension fund client. Vanguard was the first retail index fund.

[6]Vanguard enters the retail and institutional markets of the same investment style at different points in time. In the rest of this paper, I use "category" to refer to an investment style and investor type combination. For example, large-cap growth retail and large-cap growth institutional are two separate categories in the same investment style. Other companies offered index funds earlier than Vanguard in only one retail and two institutional categories, and their funds had limited impact. For example, Colonial Index Trust offered a small-cap index fund in 1986, but it carried a front load of 4.75 percent and an annual expense ratio of about 1.50 percent. The fund drew limited investor assets, and went out of business in 1993.

2.2 Exchange-Traded Funds

An exchange-trade fund (ETF) is an investment fund traded intraday on the stock exchanges at market prices. Most ETFs seek to replicate the same returns as a particular market index. The main difference between an ETF and a traditional index fund is the process of "creation/redemption in kind". In this process, an "authorized participant" (AP) acquires the securities that the ETF wants to hold in large blocks ("creation units"), exchange them with the ETF provider for ETF shares, and sells the shares in the secondary market. To redeem shares, an investor sells in the secondary market or sells the creation units back to the ETFs in exchange for the underlying securities. The APs, rather than the fund, bear the cost of trading when marginal investors purchase or redeem the fund – this is the main reason that allows the ETFs to have a lower expense ratio than the traditional index funds. In addition, the ETFs also manage to lower cost due to the elimination of back-office shareholder accounting. The first ETF in the United States—S&P Depository Receipts (SPDR)—was introduced on the American Stock Exchange in 1993. Table 1, Panel B lists the timing of ETF entry into various investment styles. At the end of 2011, the market share of ETFs among passively invested assets reached close to 35 percent (Figure 1, dashed line).

3 Theoretical Framework

This section presents a market segmentation model of the mutual fund industry to illustrate both the equilibrium among actively managed funds before the introduction of the index fund and the effect of the low-cost index fund entry.

Model Setup The model setup mirrors Ellison (2005). Two actively managed funds with the same benchmark are located at the opposite ends of a Hotelling line. Locations on the line represent degrees of active management. Fund 1 is located at 0 and is a "closet indexing" fund: While actively managed, it holds a portfolio that is similar to the benchmark. Fund 2 is located at 1 and is a truly active fund that is maximally differentiated from the benchmark. Each fund offers two vertically differentiated products: Investors can either purchase the fund alone (direct channel, let the fee be p_{iD}) or purchase the fund bundled with financial advice (broker channel, let the fee be p_{iB}). The funds are produced at marginal cost c and financial advice has zero marginal cost. Investors observe all fees. For simplicity, assume that active management does not imply different expected returns but serves to match the tastes of the investors. Therefore, the expected returns for both funds are equal.[7]

Investors differ on two dimensions. First, they exhibit different fee sensitivity. Let the sensitivity be γ. Consider three types, L, M and H, with $0 < \gamma_L < \gamma_M < \gamma_H \leq 1$, and a unit mass of investors in each type. Second, the investors have exogenous tastes $\theta \sim U[0, 1]$ for the degree of active management.[8] γ and θ are independent. An investor purchases zero or one unit from one fund in the market. An investor with type (γ, θ) has the following utility:

[7]This assumption does not drive the results of the model.

[8]For example, an investor with high θ can have a strong preference for having the *possibility* to outperform the benchmark, whereas an investor with low θ may be risk-averse toward falling below the benchmark.

$$u(\gamma,\theta) = \begin{cases} v - \theta - \gamma p_{1D} & \text{if} \quad 1D \\ v + w - \theta - \gamma p_{1B} & \text{if} \quad 1B \\ v - (1-\theta) - \gamma p_{2D} & \text{if} \quad 2D \\ v + w - (1-\theta) - \gamma p_{2B} & \text{if} \quad 2B \\ 0 & \text{if} \quad \text{None,} \end{cases}$$

where v is the utility from investing in a portfolio that exactly matches the investor's taste and w is the utility of financial advice. $\gamma \leq 1$ reflects the evidence in the literature that retail investors do not fully take fee information into consideration, either due to oversight (Barber et al. (2005)) or due to low financial literacy (Choi et al. (2010)). Note that an investor with lower γ has a higher valuation for the financial advice in terms of dollars ($\frac{w}{\gamma}$), and, all else equal, is more likely to choose the broker channel over the direct channel.[9]

Proposition 1. Equilibrium before Index Fund. Under the parameter restrictions discussed in Appendix 1, there exists a symmetric equilibrium where L and M types purchase through the broker channel and H type purchases through the direct channel. Equilibrium prices $p^*_{1D} = p^*_{2D} = c + \frac{1}{\gamma_H}$ and $p^*_{1B} = p^*_{2B} = c + \frac{1}{\bar{\gamma}}$, where $\bar{\gamma} = \frac{\gamma_M + \gamma_L}{2}$.

Sketch of proof: Figure 2 illustrates the division of the market in the equilibrium before

[9] One can write a model in which investors have heterogeneous valuation for financial advice, which can generate the same predictions through a similar selection mechanism. However, as presented later in the paper, I do find direct evidence on heterogeneous fee sensitivity. An alternative interpretation for γ is the marginal utility of income. In other words, wealthier investors, with lower marginal utility of income, are more likely to select into the broker channel. The interpretations of γ do not affect the results, though it may affect the policy implications. I do not have micro-level data on the characteristics of individual investors to distinguish among the explanations. However, ICI (2004) survey results show that investors in the broker channel have slightly lower median income and lower median financial assets than those in the direct channel, thus it is not very likely that γ represents the marginal utility of income.

the index fund entry. Fund i solves the following problem for profit maximization:

$$\max_{p_{iD},p_{iB}} [\frac{1}{2} - \frac{\gamma_H}{2}(p_{iD} - p^*_D)] \cdot (p_{iD} - c) + [1 - \frac{\gamma_M + \gamma_L}{2}(p_{iB} - p^*_B)] \cdot (p_{iB} - c),$$

which gives

$$p^*_{1D} = p^*_{2D} = c + \frac{1}{\gamma_H}$$

and

$$p^*_{1B} = p^*_{2B} = c + \frac{1}{\bar{\gamma}}.$$

The mark-up in each channel is determined by the inverse fee sensitivity of the average investor in that channel; thus, having the marginal type M in the advisory channel *benefits* the L type. This is an equilibrium when the investors' incentive-compatibility (IC) constraints are satisfied, and when funds cannot profitably deviate to another pricing strategy. Appendix 1 gives the conditions under which these constraints are satisfied.

Entry of Index Fund The objective of an index fund is to track the benchmark. Therefore, the index fund is located at 0. The index fund delivers the same expected pre-fee return as both Fund 1 and Fund 2, thus the utility from investing in one unit of index fund for investor (γ, θ) is $u_I = v - \theta - \gamma p_I$. An index fund is produced at a lower marginal cost δc, where $\delta \in [0, 1)$. The index fund is sold only through the direct channel.[10]

Proposition 2. **New Equilibrium.** Under parameter restrictions described in Appendix 2, there exists a new equilibrium. Fund 1 offers only in the broker channel and resets

[10]This assumption is to capture the fact that Vanguard index funds are distributed through the direct channel. Broker-sold index funds became available (and relatively popular) later in the sample period. The model result can still hold even when we assume that index funds enter through both channels.

its fee to $p^{*'}_{1B} = c + \frac{1}{\gamma_L}$. Fund 2 offers in both the direct channel and the broker channel. Its fees $p^{*'}_{2D} = \frac{\delta+2}{3}c + \frac{1}{\tilde{\gamma}}$ where $\tilde{\gamma} \equiv \frac{\gamma_H+\gamma_M}{2}$, and $p^{*'}_{2B} = c + \frac{1}{\gamma_L}$. The index fund fee $p^*_I = \frac{2\delta+1}{3}c + \frac{1}{\tilde{\gamma}}$.

Sketch of proof: All existing investors in 1D switch to the index fund, because the latter is a perfect substitute at a lower price. Now, the IC constraint for the M type in 1B is likely to be violated: M type switches to the index fund iff the fee difference between 1B and the index fund exceeds the M type's valuation for financial advice – i.e., iff $\gamma(p'_{1B} - p_I) > w$. Conjecturing that in the new equilibrium, competition lowers p_{2D}, the IC constraint for the M type in 2B is also likely to be violated. As Appendix 2 discusses, when the parameters satisfy certain restriction (roughly speaking, when the M type is fee-sensitive enough), the funds retain only the L type in the broker channel in the new equilibrium.

Figure 3 illustrates the equilibrium after index fund entry. In the direct channel, Fund 2 solves

$$\max_{p'_{2D}} [1 - \tilde{\gamma}(p'_{2D} - p^*_I)] \cdot (p'_{2D} - c),$$

and the index fund solves

$$\max_{p_I} [1 + \tilde{\gamma}(p^{*'}_{2D} - p_I)] \cdot (p_I - \delta c),$$

which gives

$$p^{*'}_{2D} = \frac{\delta+2}{3}c + \frac{1}{\tilde{\gamma}}$$

and

$$p^*_I = \frac{2\delta+1}{3}c + \frac{1}{\tilde{\gamma}}.$$

In the broker channel, Fund i solves

$$\max_{p'_{iB}}[\frac{1}{2} - \frac{\gamma_L}{2}(p'_{iB} - p^{*'}_{B})] \cdot (p'_{iB} - c),$$

which gives

$$p^{*'}_{1B} = p^{*'}_{2B} = c + \frac{1}{\gamma_L}.$$

Because the M type selects out of the broker channel, competition between 1B and 2B in fact softens, resulting in a fee increase in the broker-channel. This is the selection effect which is the key result of the model.[11] Two countervailing forces affect the price in the direct channel. On the one hand, due to the inclusion of the M type, the average investor in 2D has lower price sensitivity. On the other hand, competing with the low-cost index fund puts downward pressure on price. When the cost saving of the index fund $1 - \delta$ is large enough, the price competition effect dominates.[12]

Extensions The model can potentially be extended to allow the funds to choose their locations on the Hotelling line. d'Aspremont et al. (1979) shows that the equilibrium follows "*maximal* differentiation".[13] Applying this principle, in the equilibrium post index fund

[11]The model and the empirical analysis do not distinguish between two cases; one where the M type leaves the broker channel even holding the broker-channel fee constant, and one where the M type leaves *because* of the broker channel fee increase. Both can be called the selection effect, and both make the same empirical predictions. The empirical analysis makes sure the fee increase is *not* driven mechanically by the fund outflow.

[12]Another model in which prices can go up with competition is Gabaix et al. (2013) which shows that when consumers' preferences for a good is influenced by random noise, prices may increase or decrease with competition, depending on the tail properties of the noise distribution. To use this theory to explain the divergent price reactions, the distributions of the noise terms need to have specific and different properties across the channels. In contrast, the reasoning behind the selection mechanism does not require assumptions on distributional properties.

[13]The authors show that when we replace the linear transportation cost in the original Hotelling model with quadratic cost, the two firms will hit corner solutions in their location choice problem and locate at the

entry, Fund 1 has an incentive to shift the portfolio away from the benchmark, in order to soften the price competition with the index fund.

The model can also be adapted to the study of ETF entry into the index fund market. Despite the homogeneity in portfolios, index funds are differentiated by non-portfolio features (Hortacsu and Syverson (2004)),[14] such as brand preference or trust (Gennaioli et al. (2012)). We can think of θ as a parameter for any form of general non-portfolio differentiation. Figure 4 represents the equilibrium after the ETF entry. ETFs are traded directly on stock exchanges, thus the model considers them as sold through the direct channel. The predictions are similar to those in Proposition 2, except that due to their identical portfolios, ETFs and the directly sold index funds can engage in Bertrand competition.[15]

4 Data and Summary Statistics

4.1 Data

The main data set used is the CRSP Survivor Bias Free Mutual Fund Database, and the sample is limited to general U.S. domestic equity mutual funds. Appendix 3 details the

maximal distance from the competitor. With linear transportation cost, however, there is no pure strategy equilibrium when the two firms are located close to the center of the line. Economides (1984) shows that the maximal differentiation principle also applies to a circular competition model. Although it is possible to extend the model with quadratic cost functions and a stage zero game in which the funds choose locations for their portfolios, the model becomes less tractable with quadratic costs, and it is not the focus of this paper to solve the extended model.

[14]Hortacsu and Syverson (2004) estimate a search cost model and argue that horizontal differentiation cannot be separately identified from search cost.

[15]The "price war" between index funds and ETFs has been frequently featured in the news. For example, an Advantage share of the Fidelity Spartan 500 Index Fund today costs five basis points, and so does an Admiral share of Vanguard 500 Index Fund. In comparison, the Schwab U.S. Broad Market ETF costs four basis points, and the Vanguard S&P 500 ETF and the iShares Core S&P 500 ETF cost five basis points and seven basis points, respectively (Ludwig (2012)). As Burton (2011) describes, the expense ratios are "slashed to bone".

procedure of sample selection, data cleaning and variable construction. Active share and tracking error data are from Antti Petajisto's website (Petajisto (2013)). The monthly benchmark index returns are from Compustat North America Index Prices data set. The distribution channels of retail funds are approximated using distribution fee items with criteria according to ICI (2013). A retail fund share with no front or back load, and with an annual distribution fee of no more than 25 bps, is classified as direct-sold. A retail fund share with front load, back load, or an annual distribution fee of more than 25 bps is classified as broker-sold.

4.2 Summary Statistics

The mutual fund market has grown considerably over the period under study. Table 2 presents the size of the active fund and index fund markets in the sample across the broker, direct and institutional channels, reported as three snapshots in 1970, 1990 and 2005. Among active funds, more than half of the retail assets are in the broker channel. The institutional channel is small, but it has expanded substantially.[16] The market share of index funds (not including ETFs) grew from three percent of the open-end mutual fund market in 1990 to 12 percent in 2005.[17]

Table 3 presents the summary statistics of the fund characteristics across the distribu-

[16]Bergstresser et al. (2009)and Del Guercio and Reuter (2013) use distribution channel data from the Financial Research Corporation (FRC). The FRC data are not sufficient to cover the entire study period in this paper. However, the summary statistics here are similar to those in the above papers. The empirical investigation includes the data on the institutional channel and allows their reactions to competition to be different from those in the retail direct- and broker-channels. All results in this paper are robust to dropping the data on institutional share classes.

[17]When the index funds initially entered the market, they were sold primarily through the direct channel. In 1990, less than three percent of the assets in index funds were invested through the broker channel. By 2005, the broker channel share of index fund assets had grown to 18 percent. The fact that the index fund market also evolved into a segmented market gives rise to the ETF study, which examines the effect of the ETF entry on index funds in the direct channel and the broker channel.

tion channels over the sample period and, for comparison, also reports the characteristics of Vanguard index funds. Total fee is calculated as the annual expense ratio plus amortized loads. The loads (including front load and back load) are amortized over seven years, which, according to Sirri and Tufano (1998), was the average holding period of retail investors at that time.[18] Several patterns emerge from the table, highlighting index funds as a large-scale and low-cost competitor for the incumbent actively managed funds. An average Vanguard index fund charges 21 basis points of assets under management, compared with 2.22 percent and 1.16 percent for actively managed funds in the broker and the direct channels, respectively. Moreover, on average, a Vanguard retail fund is more than ten times larger than the incumbents it competes with. The difference in the portfolios between the actively managed funds and index funds is also pronounced, as seen in various measures, including the turnover ratio, defined as the fraction of a fund's holdings that have changed over a year, the active share, which is the sum of portfolio weight differences between an actively managed fund and its benchmark index (Cremers and Petajisto (2009)), and the R-squared from regressing the monthly portfolio returns on the underlying benchmark returns.

Table 4 shows the summary statistics by investment styles. We observe that, on average, funds investing in large-cap stocks are larger and less costly than funds investing in small-cap stocks. Small-cap funds are also more actively managed than large-cap funds, as seen in the higher active share and the lower R-squared over the benchmark. Additionally, growth styles have higher turnover ratios and charge higher fees than value styles do.

5 Empirical Framework

The main analysis adopts a difference-in-differences specification in the form of:

[18]The results are robust to using alternative amortization periods, as shown in Table 9.

$$
\begin{aligned}
p_{ikjt} = \alpha_i + \tau_0 PostVG_{jt} + \tau_1 PostVG_{jt} \times Direct_{ik} + \tau_2 PostVG_{jt} \times Inst_{ik} \\
+ \lambda_1 Direct_{ik} + \lambda_2 Inst_{ik} + X^{'}_{ikjt}\delta + \phi_j + \psi_t + \eta_j \times t + \epsilon_{ikjt},
\end{aligned} \tag{1}
$$

where p_{ikjt} is the total fee of share class k of fund i in category j at time t. $PostVG_{jt}$ indicates the presence of Vanguard index fund(s) in category j at time t. $Direct_{ik}$ and $Inst_{ik}$ are indicators for whether share class k of fund i is in the direct channel or the institutional channel. The coefficient τ_0 gives the effect in the broker channel. The net effect for the direct channel is given by $\tau_0 + \tau_1$ and for the institutional channel by $\tau_0 + \tau_2$. The control variables in X_{ikjt} include lagged log size of the share class, lagged log size of the fund, share class age in years, lagged net market-adjusted return, lagged turnover ratio, and lagged log fund family size.[19] ϕ_j represents category fixed effects to control for mean differences in fees across categories. $\eta_j \times t$ allows the fees across categories to be on separate linear trends, ψ_t, the year fixed effects, control for the levels of average fees over time. A set of fund fixed effects α_i is included to control for unobserved but fixed fund characteristics, such as fund manager abilities. The sample is limited to the categories that Vanguard entered during the sample period 1970-2005.[20]

The specification first takes a difference in the fees of actively managed funds before and after the entry of the Vanguard index fund. It then compares the differences in the

[19] Some of the control variables can be affected by the entry event, most notably fund size. However, including size allows me to control for a direct effect of size change on fees. As the robustness section shows, the results are robust to not including control variables.

[20] The categories that Vanguard did *not* enter during this period were nonetheless affected by the competitive pressure from passive funds over time, most notably by the introduction of the iShares ETFs in 2000. Index fund companies other than Vanguard had also offered other products along the way, but it is difficult to pin down which introduction had significant effect. For these reasons, the categories not entered by Vanguard are excluded from the difference-in-differences analysis. I exploit observations in the excluded categories for the falsification tests.

treatment and control groups and takes a second difference. At each time t, the categories not currently affected by entry serve as a control group for the categories which Vanguard is entering. Lastly, the regression allows the treatment effects to be different for the share classes in the broker, direct, and institutional channels.

Because the series of staggered entry spreads out over more than two decades, I restrict the estimates to five-year windows before and after each entry. To achieve this, the $PostVG$ dummy in Equation (1) is split into $PostVG4$ (an indicator that equals 1 for the years 0-4 post Vanguard entry in the styles it enters) and $PostVG5+$ (an indicator for year 5 and later post Vanguard entry), to estimate short-term and long-term effects separately. The reason is that the long-term effects are more likely to be affected by confounding economic forces. Similarly, to shorten the pre-period window, the specification includes $PreVG6-$ (an indicator for six years before Vanguard entry and earlier). In addition, as a Granger-type test for pre-existing trends, $PreVG1$ is included to indicate the one year before Vanguard entry. All of the pre- and post-treatment dummy variables are also interacted with the channel indicators. The full specification is given by:

$$\begin{aligned}
p_{ikjt} = \alpha_i &+ \tau_0 PostVG4_{jt} + \tau_1 PostVG4_{jt} \times Direct_{ik} + \tau_2 PostVG4_{jt} \times Inst_{ik} \\
&+ \tau_3 PostVG5+_{jt} + \tau_4 PostVG5+_{jt} \times Direct_{ik} + \tau_5 PostVG5+_{jt} \times Inst_{ik} \\
&+ \tau_6 PreVG1_{jt} + \tau_7 PreVG1_{jt} \times Direct_{ik} + \tau_8 PreVG1_{jt} \times Inst_{ik} \\
&+ \tau_9 PreVG6-_{jt} + \tau_{10} PreVG6-_{jt} \times Direct_{ik} + \tau_{11} PreVG6-_{jt} \times Inst_{ik} \\
&+ \lambda_1 Direct_{ik} + \lambda_2 Inst_{ik} + X_{ikjt}\delta + \phi_j + \psi_t + \eta_j \times t + \epsilon_{ikjt}, \qquad (2)
\end{aligned}$$

Identification Assumption For equation (2) to identify the causal effects of index fund competition, we must assume that the pre- and post-treatment dummy variables and their interactions with the channels are uncorrelated with the error term ϵ_{ikjt}. In a panel regression setup after controlling for market fixed effects and fund fixed effects, this identification assumption is satisfied if the exact *timing* of the entry is uncorrelated with any omitted variables that determine fund fees.

The identification assumption is supported by the following three pieces of evidence. First, a Cox proportional hazard model examines the determinants for the timing of the entry using various observable market characteristics. There may be a concern, for example, if Vanguard chooses to enter the investment styles that have done well or poorly in the recent past. The result in Appendix Table A1 shows that this is not the case: Vanguard entry is not correlated with lagged fees, returns, fund flows, concentration, or other market characteristics. Second, to explain what drives the entry timing, anecdotal evidence from Bogle (1997) is summarized in Table 1, Panel A. As we can see, most decisions to introduce a Vanguard index fund are a function of *idiosyncratic* considerations by Vanguard CEO John Bogle. Third, the Granger test shows no pre-existing trends in the outcome variables *prior* to Vanguard entry. For example, one potential concern is that the fees in the direct channel were already declining, and Vanguard entered to be part of the trend. The Granger test suggests that such reverse causality is unlikely.

6 Results

The model in Section 3 implies the following hypotheses about the effects of the competition from index funds:

1. Price competition: To the extent that the cost advantage of the index fund relative to active funds is large enough, active funds in the direct channel lower fees.

2. Selection: Active funds in the broker channel increase fees.

3. The entry of index fund negatively affects the fund flows to actively managed funds in the broker channel.The effects on fund flows to the direct channel can go in either direction.[21]

4. Fund flows in the broker channel are less sensitive to fees than fund flows in the direct channel, conditional on past performance.

5. In response to index fund entry, actively managed funds engage in product differentiation by increasing the level of active management. The increase should be larger for ex ante closet-indexing funds.

6. In response to ETF entry, traditional index funds in the direct channel lower fees, while traditional index funds in the broker channel increase fees.

This section first presents the main empirical result of this paper and its robustness. Next, it presents the results from testing the additional predictions of the model.

6.1 Reactions in Fees

Divergent Fee Reactions Table 5 presents the estimates from the basic specification in Equation (2). Column (1), which is the baseline result, highlights the divergence in fee reactions across mutual fund distribution channels. The entry of Vanguard index funds *increases* the fees in the broker channel by 10.5 bps, while, on net, it *decreases* the fees in the direct channel by 7.4 bps (p-value<0.10). Relative to the mean fee levels before entry (2.13 percent and 1.24 percent, for the broker channel and the direct channel respectively), the effects translate into a five-percent fee increase in the broker channel, and a six-percent decrease in the direct channel. The coefficients on the covariates are as expected: Fees are negatively correlated with fund size, age, and fund family size, due to decreasing returns

[21] The direct channel is negatively affected by the outflows to the index funds. Meanwhile, the direct channel absorbs part of the fund flow leaving the broker channel.

to scale, and positively correlated with the turnover ratio. Funds with *better* market adjusted net performance charge *lower* fees, a result consistent with Gil-Bazo and Ruiz-Verdú (2009). Columns (2) and (3) show that the results are not driven by including the data on institutional shares or the observations beyond five years post Vanguard entry. Column (4) demonstrates that the results are robust to keeping only the funds that existed before Vanguard entry. Column (5) keeps a balanced sample of the funds that existed for at least three years before and at least five years after Vanguard entry – conditional on surviving five years, the estimate of the fee decrease in the direct channel becomes 15 bps (a 12-percent decrease from mean pre-entry fee, p-value<0.01), suggesting that the fund shares that exited within five years had higher fees. Columns (6) and (7) re-estimate equation (2) using data at the fund-channel level[22] and including fund-channel fixed effects. This estimator exploits only within-fund-channel variations and measures the effects of index fund competition on *changes* in fees. The dependent variable in Column (6) is the equal-weighted average fee for a fund-channel, and the estimate changes little compared with Column (1). The dependent variable in Column (7) is the value-weighted average fee, and the result is slightly smaller, implying that more actions on fee adjustments come from smaller share classes.

Because mutual fund fees are serially correlated and are cross-sectionally correlated within fund families, the standard errors are clustered at the fund family level. Ideally, we want to cluster the standard errors by category; however, the number of categories is small in the present study, which can bias the standard error estimate (Angrist and Pischke (2008)). The robustness section below shows that the results are robust to clustering at the category level, to two-way clustering by both category and year, as well as to using group means, which

[22]For fund-channel-level data, I combine the observations on share classes within the same fund and the same channel. Size is calculated as the sum of size of share classes within the fund-channel. Age is the maximum share class age for the fund-channel. Other covariates are size-weighted averages for the fund-channel.

works well even when the number of groups is small (Donald and Lang (2007)).

It is important to note here that an analysis of the Granger test shows no pre-existing differential trends in the fees in the one year prior to Vanguard entry. The coefficients on $PreVG1$ and its interactions with the channels are small and insignificant. This test reassures that there are no trends in fees across channels before the entry of Vanguard index funds, thus the estimated fee changes can be seen as *caused* by the entry.

As the effects in the two channels go in opposite directions, a back-of-the-envelope calculation shows that the effect on aggregate fee is small. Following Table 5, Column (6), an average direct-sold fund decreases its fee from 1.24 percent to 1.13 percent, whereas an average broker-sold fund increases its fee from 2.13 percent to 2.22 percent. Taking the market share of the broker channel to be 53 percent (Bergstresser et al. (2009)), the net change in weighted-average aggregate fee is minus 0.2 percent. However, the magnitudes of the effects in the respective channels are economically significant. Consider an investor in the broker channel paying 2.13 percent per year on an investment of $100K, and the fund earns a five-percent pre-fee return annually. Over a 30-year period, the investment grows to $227K. A fee increase to 2.22 percent implies that the investment grows to $219K, so the investor loses 2.7 percent of her wealth due to the increase in the broker channel fee. In contrast, an investor in the direct channel ex ante pays 1.24 percent annually, and her investment grows to $297K over 30 years. A fee decrease to 1.13 percent allows her to grow her investment to $307K – the investor saves 3.3 percent of her wealth on active investment due to the decrease in direct channel fees after the index fund entry.[23]

[23] Described here is the welfare gain and loss caused by competition from the index fund entry for investors remaining in actively managed funds. Of course, the index fund also offers a lower-fee option in the direct channel. The welfare gain for the investors who switch to choose the index fund is larger.

Divergent Reactions Not Driven by Fund Characteristics Funds across distribution channels, as Table 3 shows, exhibit different characteristics. For example, share classes in the broker channel are smaller than those in the direct channel. To account for the possibility that different characteristics can affect both the funds' channel choices and the responses to competition, the regressions in Table 6 allow the covariates to interact with the pre-, post-treatment indicators. Column (1) controls for lagged log sizes of both fund i and share class ik interacted with the $PreVG6-$, $PostVG4$, $PostVG5+$ dummy variables. The rationale is that large and small funds have different cost curves which may affect their positions to react to competition. Column (2) controls for all covariates interacted with the pre- and post-Vanguard dummy variables. We see that the estimated *differential* response between the broker channel and the direct channel (as well as between the broker channel and the institutional channel) is unchanged, suggesting that none of the observable fund characteristics drive the divergent responses across the channels.

The bottom panel of Table 6 calculates the net treatment effects by channel. Because of the inclusion of the interaction terms between the shock and the covariates, the net effects can no longer be directly computed. Instead, the effect on a channel is evaluated at the means of the covariates for this channel. The net effects are similar to those obtained from the basic specification.

Dynamics in Effects The results presented in Table 5 and Table 6 give the average effects estimated over a five-year window. To explore the speed at which the funds respond, as well as the persistence of the effects, I run a specification that is similar to equation (2) but augmented with indicators for the lead and lag years relative to the entry. Table 7 reports the coefficients on the treatment effect dynamics, where the observations are at the fund-channel level and fund-channel fixed effects are controlled for. Figure 5 gives a visualization of the timing of the effects. The estimated effect in the broker channel becomes positive in the first year of entry. It grows to a statistically significant nine basis points in year 2 and stabilizes at that level thereafter. The differential reactions in the direct and the institutional channels become significant within one year after the entry. The net effects reach about -14 bps for the direct channel and -12 bps for the institutional channel by year 4. The competition from index funds have long-lasting effects on the pricing strategies of actively managed funds: The estimated effects persist beyond five years.

Falsification Test Although we have strong reason to believe that the timing of Vanguard entry is orthogonal to the incumbents' fee policies, there may still be concerns that the entry is correlated with some trends in fees across the channels. Two falsification tests are constructed to rule out such possibility.

The first falsification test uses data on the four equity fund styles that are *not* entered by Vanguard during the sample period 1970-2005: Mid-Cap Growth, Mid-Cap Value, Multi-Cap Growth, and Multi-Cap Value. The fictional entry years for Mid-Cap (Multi-Cap) Growth/Value are assigned to be the actual entry years for related styles – i.e., Mid-Cap (Multi-Cap) Core. The estimated responses to fictional index fund entry should be small or zero. The second falsification test uses the same sample as the main regression sample, but with *random* entry years for the styles. Again, the estimated effects of random entry should be none. Table A2 in the Appendix reports the results of the falsification tests: Both falsification tests support that the baseline findings are indeed caused by the actual entry of Vanguard index funds, rather than by spurious correlations in the data.

Robustness Table 9 shows the results from a list of additional robustness checks. Columns (1)-(2) investigate alternative ways of clustering the standard errors, and the results are robust to clustering the standard errors by category, as well as to two-way clustering by category and year. Columns (3)-(4) show that not including the covariates or the linear trends by categories does not change the result. Columns (5)-(7) investigate the estimates using alternative dependent variables. Because the average fund-holding period (seven years) is a proxy, here the analysis shows that the results are robust to using other assumptions of holding periods or using only the annual expense ratio as the dependent variable. Column (8) includes category-by-year fixed effects to fully control for category-specific shocks that may correlate with Vanguard entry, for example, the popularity of the investment strategy. The estimate exploits only the variation across the broker and direct channels in a given category-year, and the estimated differential reaction between the channels remains unchanged. Column (9) controls for channel-by-year fixed effects. This specification takes away much variation in the post-treatment variables interacted with the channel indicators, so the results are much weaker. The fee increase in the broker channel remains significant only at 0.10 level.

Table 10 estimates the effects of the low-cost index fund entry on aggregate fees of actively managed funds. The data are grouped at the investment style-channel-year level, and the standard errors are clustered by category. Column (1) uses the equal-weighted average fee as the dependent variable. The analysis serves two purposes. First, it is an alternative method of addressing clustered standard errors; and second, it scrutinizes the effects of outliers. Column (2) uses the asset-weighted average fee, which allows us to assess whether the fee changes are economically important – in other words, they are not driven by small funds only. Column (3) is estimated using a two-step procedure that adjusts for micro covariates. The first step is to regress share-class-level fees on micro covariates and group dummy variables to estimate the group effects, and the second step is to use the estimated group effects as the dependent variable in a group-level regression. The divergence in the fee reactions is robust across all specifications using aggregate data.

6.2 Asset Attrition

The model predicts that actively managed funds, especially assets in the broker channel, experience attrition as a result of the selection mechanism. To examine this hypothesis, the effects on net flow percentages for actively managed funds are estimated using data at the fund-channel level and at the fund level. Following the literature, net flow percentages are calculated as

$$Flow_t = \frac{TNA_t - TNA_{t-1} \cdot (1 + r_t)}{TNA_{t-1}}, \tag{3}$$

where TNA_t is the total net asset at time t, and r_t is the net-of-fee return from $t-1$ to t.

Table 11 reports the estimates. Columns (1)-(2) show that the flow rates to fund assets in the broker channel decrease by three to five percent. The reduction in the flow rates for the direct channel is about six percent, although the difference between the two retail channels is not statistically significant. The reduction in the flow rate for the institutional channel is about nine percent. Columns (3)-(4) suggest that the average reduction in the flow rate to a retail actively managed fund is five to six percent.

Next, I examine the effects on the size and the fee revenue collected (fee multiplied by assets under management) by the actively managed funds, using data at the fund-channel level. Column (5) suggests that the asset size of funds in the broker channel is negatively affected, although the effect is not statistically significant. Column (6) indicates that the revenue in the broker channel change little, which is reasonable given that it is jointly affected by a size reduction and a fee increase. Columns (5)-(6) also show, in contrast, that the funds in the direct channel experience a nine-percent reduction in size (p-value$<$0.01) and a 15-percent reduction in revenue (p-value$<$0.05).

6.3 Cross-Sectional Heterogeneity

The analysis explores the cross-sectional heterogeneity in the effects, which reflects the nature of index funds' threat for actively managed funds. Using data on the funds that existed before the index fund entry, I divide the sample into large and small fund families, as well as funds with high and low active share, according to the ex ante values. The idea is that smaller fund families may offer more specialized funds that do not compete directly with the index fund, and the funds with higher active share are also affected less. Table 12 Columns (1)-(2) suggest that the fee reaction, especially the fee cut in the direct channel, is stronger in larger fund families than in smaller families (although the difference in the estimates between the subsamples is not significant). The results on fund flows in Columns (3)-(4) support the idea that larger fund families are more negatively affected–they experience larger fund outflow, especially from the direct channel (the difference has p-value$<$0.05). Consistent with the index funds being a closer substitute for closet indexers, Columns (5)-(8) indicate that the funds with ex ante below-median active share are affected more negatively by the index fund entry. The outflows they experience are larger, both in the broker channel and in the direct channel (p-values are 0.04 and 0.14, respectively), compared with the ex ante truly active funds. The adjustments in fees are not significantly different between the two subgroups. As we will see in Section 6.6, the reason lies in the portfolio adjustments undertaken by the closet indexers.

The analysis then examines whether the fee reactions differ for funds with high and low past performance. If the fees charged by active fund managers reflect their abilities, the funds with higher performance relative to the benchmark should be affected less by the index fund entry. The results in Columns (9)-(12) do not support this notion. In fact, there are no different reactions between the two subsamples with high and low performance.

6.4 Alternative Explanations

There may be alternative arguments for why fees in the broker channel increase. For example, one explanation is that fund companies may face higher advertising expenditures, or may have to compensate their brokers more. Examining the components of the fee changes can shed light on this alternative. Following Gil-Bazo and Ruiz-Verdú (2009), I divide the total fee into a marketing fee (amortized loads plus annual 12b-1 fee) and a non-marketing fee, or mostly management fee (annual expense ratio minus 12b-1 fee). If any change in advertising cost or in broker compensation drives the result, we expect to see the effects come solely from marketing fees. However, the results in Table 8 show that about half of the fee increase (5.1 bps out of 10.6 bps) in the broker channel comes from an increase in the marketing fee, while the other half (5.5 bps) comes from an increase in the non-marketing fee.[24]

Another alternative explanation is that the broker-sold funds shrink, and fees can go up due to decreasing returns to scale. However, the magnitude of the fee increase exceeds what can be explained by the change in size alone. First, all estimates of the effects are already conditional on size. Second, the coefficients on lagged log size in Table 5 suggest that fees

[24]The focus of this paper is on investor heterogeneity, and the analysis examines the total fees *paid* by investors, instead of how the fee revenue is split between the fund company and the broker entity. The revenue sharing arrangement is a separate issue, not covered by this article but interesting for future studies. Another potential concern is that a subset of the fund companies charge management fees that decline mechanically as assets under management rise. In the CRSP database, management fees become available after 1998; thus, it is not feasible to estimate the effects on the management fee using the index fund entry shocks. However, in Table A3 in the Appendix, I estimate the effects of ETF entry on different components of fees, and the effects are not driven by the management fee.

decline slowly with fund size. Given that funds in the broker channel shrink by about six percent (Table 11, Column (5)), this corresponds to a 0.3-basis-point increase in fee, far below the estimated fee increase of nine basis points in the broker channel.

6.5 Flow-Fee-Sensitivity across Channels

The model hinges on heterogeneous investor fee sensitivity, and a direct implication is that fund flows in the direct channel should be more sensitive to fees than fund flows in the broker channel. As a test of this hypothesis, the study estimates the cross-sectional flow-to-fee sensitivity across the distribution channels with the following regression. The model resembles that in Del Guercio and Reuter (2013) for flow-to-performance sensitivity estimation.

$$\begin{aligned} Flow_{ikjt} &= \gamma_1 TotalFee_{ikjt-1} + \gamma_2 TotalFee_{ikjt-1} \times Direct_{ik} + \gamma_3 TotalFee_{ikjt-1} \times Inst_{ik} \\ &+ X_{ikjt}\delta_1 + Direct_{ik} \times X_{ikjt}\delta_2 + Inst_{ik} \times X_{ikjt}\delta_3 + \nu_{kjt} + \epsilon_{ikjt}, \end{aligned} \tag{4}$$

where $Flow_{ikjt}$ is calculated according to (3). The regression controls for covariates X_{ikjt}, including, most importantly past performance, as well as fund size, age and fund family size. The effects of the covariates are allowed to be different across the channels. The idea is to test whether fund flows demonstrate different correlations with fees, conditional on their different correlations with performance. Note that style-by-channel-by-year fixed effects ν_{kjt} control for the average net flow percentage for each style-channel-year. γ_1 measures the cross-sectional flow-fee sensitivity in the broker channel. $\gamma_1 + \gamma_2$ measures the sensitivity in the direct channel, and $\gamma_1 + \gamma_3$ in the institutional channel. The model predicts γ_2 to be negative and significantly different from zero.

Table 13 Columns (1)-(5) report the estimates following equation (4). The results across various specifications show that fund flows in the direct channel are significantly more sensitive to fees than the fund flows in the broker channel. In the cross-section of the broker channel, a one-standard-deviation increase in the total fee (60 bps) corresponds to a 2.6-percent decrease in the annual fund growth rate, whereas in the direct channel, a one-standard-deviation increase in the total fee (65 bps) corresponds to a 5.3-percent decrease. The estimates are robust to using the pre-fee return as the performance measure, including the net 4-factor alpha in addition to net market-adjusted return, or using the expense ratio as the fee measure. Column (6) uses data at the fund-channel level and controls for fund-channel fixed effects, thus it estimates how fund flows change in response to *changes* in fees. We see that fund flows in the broker channel are *not* sensitive to changes in fees, while the fund flows in the direct channel are significantly sensitive.

Table 14 investigates the selection mechanism by examining whether the flow-fee sensitivity decreases in the broker channel in the post-entry years. In a specification building on that in Table 13, Column (6), and with subsamples broken down by the channels, fund flow rates are regressed on the pre- and post-dummy variables interacted with lagged fee. In addition, lagged fee interacted with category fixed effects and year fixed effects are included to control for the average flow-fee sensitivity across categories and over time. The regression also controls for fund-channel fixed effects and investment style-by-year fixed effects. Columns (1)-(2) show that the entry of low-cost index fund, indeed, leads to a *decrease* in the fee sensitivity in the broker channel, which is consistent with the model's prediction. Columns (3)-(6) suggest that there is no such sensitivity change in the direct or the institutional channels.[25]

[25] One caveat is that there can be an endogeneity bias because the fee changes in the post-period are endogenous to the expected flows. However, the fees adjusted quickly (Figure 5), while the change in fee sensitivity is long-lasting.

6.6 Portfolio Adjustments

To soften the price competition with the index fund, active fund managers should have an incentive to differentiate their portfolios from the benchmark portfolio. To test this hypothesis, the effect of the index fund entry on the portfolios of actively managed funds is estimated using a DID regression, and Table 15 reports the results. The main outcome variable examined is the active share measure created in Cremers and Petajisto (2009).[26] Column (1) shows that the active share increases by about 0.03, or four percent of its ex ante mean value. Using alternative measures of active management reveals similar results – tracking errors of funds increase by 0.008 (Column (2)), and R-squareds over the benchmark index decrease by 0.018 (Column (3), p-value=0.10). The magnitudes of the differentiation are small, suggesting that it is potentially costly to substantially increase the level of active management. Figure 6 plots the dynamics in the effects on active share. The effect takes place immediately after Vanguard index fund entry, peaks at year 2, and stabilizes at a slightly lower level by year 3.

Next, I explore whether the ex ante closet indexers differentiate their portfolios by more compared with the ex ante truly active funds, as the index fund is a closer substitute for the former. Columns (4), (6) and (7) show that while the true active funds increase active share by 0.012 (1.3 percent of mean, statistically insignificant), the closet indexing funds increase active share by 0.036 (4.4 percent of mean, p-value<0.01). Column (5) allows for the fund size to affect the response to competition by controlling for its interaction with the competitive shock, and the result shows that the difference in size does not drive the difference in the portfolio reactions. Overall, the results confirm that the index fund competition causes

[26] The measure is defined as $ActiveShare = \frac{1}{2}\Sigma_{i=1}^{N}|w_{fund,i} - w_{index,i}|$ where $w_{fund,i}$ is the weight of stock i in the portfolio of the fund, and $w_{index,i}$ is the weight of the same stock in the portfolio of the fund's benchmark index.

actively managed funds to take on product differentiation strategies.

6.7 Results on ETF Shock to Traditional Index Funds

The entry of ETFs into the passive fund market offers another source of exogenous variation which can be exploited to test the economic mechanism highlighted in this study. Using the staggered entry of ETFs, the analysis conducts DID analysis on fees with the sample of traditional index funds over 1990-2005. In this ETF study, identification requires that the specific timing of ETF entry be uncorrelated with the fee policies of the incumbent traditional index funds. This assumption is again plausible. First, the drivers behind the timing of ETF entry (see Table 1, Panel B) are orthogonal to factors correlated with fees: The first ETFs (SPDR 500 and Mid-Cap SPDR) were set up as experiments by the AMEX to boost trading volume, and they succeeded unexpectedly. The timing of the later ETF entry was affected by the SEC approval procedure (Gastineau (2010)), as well as by personnel incidents. Second, an indicator for the one year before ETF entry is included in the DID regression as a Granger test for causality; the finding is that there are no pre-existing differential trends in the fees leading up to entry. Table 16 exhibits the summary statistics across the channels for the incumbent traditional index funds, as well as for ETFs, and shows that ETFs enter the market as a lower-cost alternative for the existing index funds.

Table 17 presents the main results of the ETF study. The entry of ETFs into a category increases the broker channel fee by 13 basis points, or 12 percent of mean, but lowers the direct channel fee by five basis points, or 11 percent of mean. The results are robust to dropping the Vanguard index funds from the sample and to dropping the observations that are more than 3 years before or more than 4 years after the shock. Moreover, the different responses to competition across the channels are not driven by the channels' characteristics.

Taking the broker channel market share to be 12 percent (Table 2), the *net* change in weighted-average aggregate fee in the traditional index fund market is minus 5.4 percent.

7 Conclusion

The majority of mutual fund studies have focused on fund performance as a mechanism to allocate resources in this important industry, yet mutual fund fees – the price mechanism – have received less attention. It has remained an open question how fees are set in this market, and how they strategically adjust in response to competition. The introductions of low-cost and large-scale index funds and ETFs – arguably the most significant competitive shocks to the landscape of the money management industry in the past few decades – provide a natural experiment to study these questions. While a simple model predicts that competition reduces fees for all incumbents, the problem is more subtle when the investor population is heterogeneous. Therefore, identifying the effects of passive fund competition across market segments is an important empirical investigation.

This study specifically estimates the effects of index fund competition on mutual fund fees and investment strategies through exploiting the staggered initial entry of index funds and ETFs. The difference-in-differences analysis yields a counter-intuitive but robust finding – the entry of lower-cost alternatives leads to divergent reactions by the incumbents across the fund distribution channels. While the actively managed funds in the direct channel cut fees to compete with the index fund, those in the broker channel actually increase fees. An examination of the reactions by traditional index funds to the wave of ETF entry leads to similar findings. The effects on consumer welfare are thus very different from what is implied by a competition model with homogeneous consumers.

To explain the divergent price reactions, I build a price discrimination model to formally illustrate the effect of low-cost competition, in particular when investors have heterogeneous fee sensitivity. The model shows that segmentation has important implications for the economic welfare of competition. While the entry of low-cost financial innovations exerts a price-competition effect as expected, it also creates a selection effect that isolates the least-price-sensitive investors – this underlying investor composition change leads to a fee increase in the broker channel. The model generates additional hypotheses that guide further empirical exploration. The analysis shows that fund flows in the broker channel are less fee sensitive than those in the direct channel, and the broker channel fee sensitivity goes down post the index fund entry. Competition from low-cost index funds also reduces the fund flows to actively managed funds and leads active funds to become more actively managed.

To summarize, the analysis in this paper shows that low-cost competition in the money management industry creates a selection effect that hurts the investors in the high-price end of the market, while benefiting the investors in the low-price end of the market. This nuanced view on the welfare transfer in this industry potentially offers an explanation for the empirical puzzle of why average mutual fund fees decline only slowly over time, despite sharply rising competition from cheaper passive investment vehicles. The findings imply that understanding the determinants of investors' fee (in)sensitivity is an area that can generate important policy implications. Examining other dimensions of the responses of actively managed funds in the continuing trend toward passive asset management is another interesting area for future research.

References

Alexander, Gordon J, Jonathan D Jones, and Peter J Nigro, 1998, Mutual fund shareholders: Characteristics, investor knowledge, and sources of information, *Financial Services Review* 7.

Angrist, Joshua D, and Jörn-Steffen Pischke, 2008, Mostly harmless econometrics: An empiricist's companion .

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, 2095–2120.

Bergstresser, Daniel, John MR Chalmers, and Peter Tufano, 2009, Assessing the costs and benefits of brokers in the mutual fund industry, *Review of Financial Studies* 22, 4129–4156.

Berk, Jonathan B, and Richard C Green, 2004, Mutual fund flows and performance in rational markets, *Journal of Political Economy* 112, 1269–1295.

Berk, Jonathan B, and Jules H Van Binsbergen, 2012, Measuring managerial skill in the mutual fund industry, *Working Paper* .

Bogle, John, 1997, The first index mutual fund: a history of vanguard index trust and the vanguard index strategy .

Burton, Jonathan, 2011, The incredible shrinking fee .

Capon, Noel, Gavan J Fitzsimons, and Russ Alan Prince, 1996, An individual level analysis of the mutual fund investment decision, *Journal of Financial Services Research* 10, 59–82.

Carlin, Bruce I, 2009, Strategic price complexity in retail financial markets, *Journal of Financial Economics* 91, 278–287.

Carlin, Bruce Ian, and Gustavo Manso, 2011, Obfuscation, learning, and the evolution of investor sophistication, *Review of Financial Studies* 24, 754–785.

Célérier, Claire, and Boris Vallée, 2014, What drives financial complexity? a look into the retail market for structured products, *Working Paper* .

Chalmers, John, and Jonathan Reuter, 2013, What is the impact of financial advisors on retirement portfolio choices and outcomes?, *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, 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, 201–235.

Christoffersen, Susan EK, and David K Musto, 2002, Demand curves and the pricing of money management, *Review of Financial Studies* 15, 1499–1524.

Coates, John C, and R Glenn Hubbard, 2007, Competition in the mutual fund industry: Evidence and implications for policy, *J. Corp. Law.* 33, 151.

Cochrane, John H, 2013, Finance: Function matters, not size, *The Journal of Economic Perspectives* 27, 29–50.

Cremers, KJ Martijn, and Antti Petajisto, 2009, How active is your fund manager? a new measure that predicts performance, *Review of Financial Studies* 22, 3329–3365.

Cremers, Martijn, Miguel Ferreira, Pedro Matos, and Laura Starks, 2011, The mutual fund industry worldwide: explicit and closet indexing, fees, and performance, *Working Paper* .

d'Aspremont, Claude, J Jaskold Gabszewicz, and J-F Thisse, 1979, On hotelling's" stability in competition", *Econometrica* 1145–1150.

Del Guercio, Diane, and Jonathan Reuter, 2013, Mutual fund performance and the incentive to generate alpha, *The Journal of Finance* .

Del Guercio, Diane, Jonathan Reuter, and Paula Tkac, 2010, Demand for financial advice, broker incentives, and mutual fund market segmentation .

Donald, Stephen G, and Kevin Lang, 2007, Inference with difference-in-differences and other panel data, *The review of Economics and Statistics* 89, 221–233.

Economides, Nicholas, 1984, The principle of minimum differentiation revisited, *European Economic Review* 24, 345–368.

Ellis, Charles D, 1975, The loser's game, *Financial Analysts Journal* .

Ellison, Glenn, 2005, A model of add-on pricing, *The Quarterly Journal of Economics* 120, 585–637.

Elton, Edwin J, Martin J Gruber, and Jeffrey A Busse, 2004, Are investors rational? choices among index funds, *the Journal of Finance* 59, 261–288.

French, Kenneth R, 2008, Presidential address: The cost of active investing, *The Journal of Finance* 63, 1537–1573.

Gabaix, Xavier, and David Laibson, 2006, Shrouded attributes, consumer myopia, and information suppression in competitive markets, *The Quarterly Journal of Economics* 121, 505–540.

Gabaix, Xavier, David Laibson, Deyuan Li, Hongyi Li, Sidney Resnick, and Casper G de Vries, 2013, The impact of competition on prices with numerous firms .

Gastineau, Gary L, 2010, *The exchange-traded funds manual*, volume 186 (John Wiley & Sons).

Gennaioli, Nicola, Andrei Shleifer, and Robert W Vishny, 2012, Money doctors, Technical report, National Bureau of Economic Research.

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 & Organization* 67, 871–885.

Gil-Bazo, Javier, and Pablo Ruiz-Verdú, 2009, The relation between price and performance in the mutual fund industry, *The Journal of Finance* 64, 2153–2183.

Greenwood, Robin, and David Scharfstein, 2013, The growth of modern finance, *Journal of Economic Perspectives* .

Hortacsu, Ali, and Chad Syverson, 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, 403–456.

ICI, 2004, Profile of mutual fund shareholders .

ICI, 2013, Investment company fact book, chapter 5 .

Khorana, Ajay, and Henri Servaes, 2012, What drives market share in the mutual fund industry?, *Review of Finance* 16, 81–113.

Ludwig, Olly, 2012, Fidelity cuts fees on index funds .

Malkiel, BG, 1973, A random walk down wall street: The time-tested strategyfor successful investing.

Matsa, David A, 2011, Competition and product quality in the supermarket industry, *The Quarterly Journal of Economics* 126, 1539–1591.

Melzer, Brian, and Donald Morgan, 2012, Competition in a consumer loan market: Payday loans and overdraft credit .

Mullainathan, Sendhil, Markus Noeth, and Antoinette Schoar, 2012, The market for financial advice: An audit study, *Working Paper* .

Petajisto, Antti, 2013, Active share and mutual fund performance, *Financial Analysts Journal* 69.

Philippon, Thomas, 2012, Has the us finance industry become less efficient? on the theory and measurement of financial intermediation, *Working Paper* .

Samuelson, Paul A, 1974, Challenge to judgment, *The Journal of Portfolio Management* 1, 17–19.

Schwartz, Eduardo S, and Walter N Torous, 1989, Prepayment and the valuation of mortgage-backed securities, *The Journal of Finance* 44, 375–392.

Sirri, Erik R, and Peter Tufano, 1998, Costly search and mutual fund flows, *The Journal of Finance* 53, 1589–1622.

Wahal, Sunil, and Albert Yan Wang, 2011, Competition among mutual funds, *Journal of Financial Economics* 99, 40–59.

Wharton, 1962, Study of mutual funds, *Washington: Government Printing Office* .

Table 1: Timing of Index Fund and ETF Entry

Panel A Index Funds

Style	Investor Type	Vanguard Entry Year	Vanguard Index Fund	First Non-VG Index Fund	Year
Large-Cap Core	Retail	1976	First Index Investment Trust (1)	B.C. Ziegler	1985
Mid-Cap Core	Retail	1987	Extended Market (2)	Dreyfus	1991
Small-Cap Core	Retail	1989	Small-Cap Index Fund (3)	Colonial	1986*
Large-Cap Growth	Retail	1992	Growth Fund (4)	Padco	1994
Large-Cap Value	Retail	1992	Value Fund (5)	Waterhouse	1998
Multi-Cap Core	Retail	1992	Total Stock Market Portfolio (6)	Fidelity	1997
Small-Cap Growth	Retail	1998	Small-Cap Growth Index Fund	Barclays (ETF)	2000
Small-Cap Value	Retail	1998	Small-Cap Value Index Fund	Barclays (ETF)	2000
Large-Cap Core	Institutional	1990	Institutional Index (7)	Firstar	1989*
Mid-Cap Core	Institutional	1997	Extended Market	Munder	1998
Multi-Cap Core	Institutional	1997	Total Stock Market Portfolio	TIAA-CREF	1999
Small-Cap Core	Institutional	1997	Small-Cap Index Fund	Federated	1992*
Large-Cap Growth	Institutional	1998	Growth Fund	TIAA-CREF	2002
Large-Cap Value	Institutional	1998	Value Fund	Schwab	1999
Small-Cap Value	Institutional	1999	Small-Cap Value Index Fund	Schwab	1999
Small-Cap Growth	Institutional	2000	Small-Cap Growth Index Fund	TIAA-CREF	2002

Drivers for product timing:

(1) In response to advocacy by academics, Malkiel (1973), Samuelson (1974) and Ellis (1975).
(2) Establish concept of “Vanguard family”; extend coverage to entire U.S. stock market.
(3) "Impulsive" termination of external contract for active small-cap fund and conversion into index fund.
(4) Introduction of S&P/BARRA growth index.
(5) Introduction of S&P/BARRA value index.
(6) Decision to make it more convenient to own the total stock market.
* indicates index funds of other fund families offered before Vanguard.
Data source: CRSP Mutual Fund Database and Bogle (1997).

Panel B ETFs

Category	ETF Entry Quarter	Name of First ETF(s)	Management Company
Large Core	1993 Q1	SPDR Trust (1)	State Street
Large Growth	2000 Q2	iShares S&P/BARRA Growth	Barclays
		iShares Russell 1000 Growth	Barclays
Large Value	2000 Q2	iShares S&P/BARRA Value	Barclays
		iShares Russell 1000 Value	Barclays
Mid Core	1995 Q2	Mid-Cap SPDR Trust (2)	Bank of New York
Multi Core	2000 Q2	iShares Russell 3000 Index	Barclays
		iShares Dow US Total Market	Barclays
Multi Growth	1999 Q1	NASDAQ-100 Trust (3)	Bank of New York
Small Core	2000 Q2	iShares S&P Small-Cap 600 Index	Barclays
		iShares Russell 2000 Index	Barclays
Small Growth	2000 Q3	iShares Russell 2000 Growth	Barclays
		iShares S&P Small-Cap 600 Growth	Barclays
Small Value	2000Q3	iShares Russell 2000 Value	Barclays
		iShares S&P Small-Cap 600 Value (4)	Barclays

Drivers for product timing:

(1) SPDR S&P 500: an experiment to save AMEX from declining trading volume.

(2) Mid-Cap SPDR: experiment to cover the roll-over stocks from S&P 500.

(3) NASDAQ-100: delayed for two years by SEC because not meeting diversification standard for investment company.

(4) Barclays iShares: timing determined by random personnel incidents.

Data source: CRSP Mutual Fund Database. Interviews with professionals.

Table 2: Market Size across Distribution Channels

This table reports the aggregate total net assets in domestic equity funds by distribution channels in 1970, 1990, and 2005. The broker vs. direct channels for retail shares are classified using information on distribution fees reported in the CRSP database. A retail fund share with no front or back load, and with an annual distribution fee of no more than 25 bps, is classified as sold in the direct channel. A retail fund share with either front load or back load, or an annual distribution fee of more than 25 bps, is classified as sold in the broker channel. The number of funds is counted after aggregating different share classes of the same fund in the same channel into one observation.

		Active Funds			Index Funds	
		1970	1990	2005	1990	2005
Aggregate Size	Broker	30.5	87.7	1,282.3	0.1	54.5
($ Billion)	Direct	3.7	60.6	954.1	3.5	248.3
	Institutional	0.0	3.3	242.9	0.5	49.1
	Total	34.2	151.6	2,479.3	4.1	351.9
Number of Share Classes	Broker	218	276	3,815	2	161
	Direct	79	265	965	11	96
	Institutional	2	24	1,029	1	88
	Total	299	565	5,809	14	345
Number of Funds	Broker	218	276	1,443	2	86
	Direct	79	266	781	11	72
	Institutional	2	24	926	1	74
	Total	299	566	3,150	14	232

Table 3: Summary Statistics of Active Funds vs. Vanguard

This table reports the summary statistics of actively managed U.S. equity funds across distribution channels using data at the share class level over the sample period 1970-2005. For comparison, the statistics on Vanguard index funds are also reported. Total Fee is calculated as the annual expense ratio plus one-seventh of the sum of the front load and the back load. When multiple levels of front and back loads are reported for a share class-year, the maximal levels are used. Share Class Size is the Total Net Assets in millions. Years since Inception are calculated using the first offer dates of the share classes. Turnover is the the fraction of a fund's holdings that have changed over a year. R-squared over the benchmark is the R-squared obtained from regressing past 24-month net monthly returns on the monthly returns of the benchmark index. Active Share is from Cremers and Petajisto (2009) and is the sum of portfolio weight differences between a fund's portfolio and the benchmark portfolio. Excess Net Return over Market is the annual net raw return minus the annual return of the market. Excess Net Return over Index is the annual net raw return minus the annual return of the benchmark. Net 4-factor alpha is 12 times the intercept obtained from regressing past 24-month net monthly returns on the monthly returns of the Fama-French three factors and the momentum factor. Data limited to the final sample used in the difference-in-differences regression, which includes all actively managed funds in the investment styles that experienced Vanguard entry during 1970-2005.

	Broker			Direct			Institutional			Vanguard Retail			Vanguard Institutional		
	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
Total Fee %	2.22	2.20	0.54	1.16	1.11	0.50	1.07	1.00	0.41	0.21	0.21	0.06	0.09	0.10	0.02
Front Load %	2.42	0	2.96	0	0	0	0.17	0	0.97	0	0	0	0	0	0
Back Load %	1.48	1.00	1.87	0	0	0	0.25	0	0.78	0	0	0	0	0	0
Expense Ratio %	1.66	1.67	0.58	1.16	1.12	0.50	1.01	0.99	0.34	0.21	0.21	0.06	0.09	0.10	0.02
12b-1 Fee %	0.54	0.50	0.41	0.05	0	0.10	0.03	0	0.12	0.01	0.01	0.01	0.01	0.02	0.01
Share Class Size $ Million	381.1	35.0	2,457.0	682.8	72.1	2,832.8	224.6	62.3	533.7	8,229.3	2,173.0	17,667.7	2,721.2	788.6	5,298.0
Years since Inception	8.9	6.0	10.9	10.9	6.0	12.3	6.4	5.0	5.0	8.3	6.0	6.4	5.2	5.0	2.7
Turnover	0.85	0.67	0.84	0.90	0.65	1.09	0.87	0.71	0.75	0.24	0.19	0.20	0.27	0.23	0.21
R-Squared over Benchmark	0.85	0.88	0.12	0.84	0.87	0.13	0.86	0.88	0.11	0.97	0.99	0.04	0.96	0.99	0.05
Active Share	0.76	0.78	0.18	0.77	0.81	0.18	0.75	0.78	0.19	0.18	0.11	0.19	0.19	0.10	0.20
Excess Net Return over Market %	-1.07	-2.18	12.04	-0.24	-1.51	13.30	-0.28	-1.55	13.06	0.76	-0.21	7.12	1.62	0.40	7.63
Excess Net Return over Index %	-1.13	-1.91	9.70	-0.31	-1.27	10.80	-0.21	-1.23	10.37	-0.14	-0.06	4.40	0.16	0.15	4.83
Net 4-Factor Alpha % (Annualized)	-1.46	-1.60	6.29	-0.43	-0.66	6.94	-0.75	-0.99	6.59	-0.10	-0.23	2.54	0.22	-0.09	3.02
Observations	24,550			8,895			6612			192			87		

Table 4: Summary Statistics across Investment Styles

This table reports the means and standard deviations of the characteristics of share classes of active general U.S. domestic equity funds across the investment styles. The top part contains the eight investment styles that experienced Vanguard entry during 1970-2005. The bottom part contains the four styles that did not experience Vanguard entry during 1970-2005. Broker, Direct and Inst are the fractions of the numbers of share classes in the broker, direct and institutional channels. Total Fee is calculated as the annual expense ratio plus one-seventh of the total loads (front load plus back load). Size is the Total Net Assets of share classes in millions. Years since Inception are calculated using the first offer dates of the share classes. Active Share is the sum of portfolio weight differences between a fund's portfolio and the benchmark portfolio, as defined by Cremers and Petajisto (2009). R-squared over the benchmark is calculated by regressing the net monthly share class returns over the past 24 months on the monthly returns of the benchmark index. Turnover is the the fraction of a fund's holdings that have changed over a year.

	Observations	Broker	Direct	Inst	Total Fee %	Size ($ Million)	Years since Inception	Active Share	R-Squared over Benchmark	Turnover
Vanguard Entry during 1970-2005										
Large-Cap Core	8,293	0.62	0.23	0.15	1.73	544.1	10.5	0.63	0.89	0.77
					(0.74)	(3,359.2)	(12.5)	(0.17)	(0.11)	(0.88)
Large-Cap Growth	6,888	0.60	0.22	0.18	1.82	474.9	9.2	0.69	0.82	0.93
					(0.75)	(1,975.6)	(11.1)	(0.14)	(0.11)	(0.76)
Large-Cap Value	4,983	0.62	0.22	0.16	1.64	879.1	12.1	0.68	0.85	0.60
					(0.70)	(4,130.9)	(15.0)	(0.13)	(0.11)	(0.65)
Mid-Cap Core	2,474	0.59	0.24	0.17	1.82	317.6	7.4	0.87	0.82	1.01
					(0.74)	(869.0)	(7.1)	(0.14)	(0.13)	(1.03)
Multi-Cap Core	6,226	0.65	0.23	0.12	1.76	316.4	9.2	0.76	0.85	0.84
					(0.73)	(1,227.7)	(10.8)	(0.16)	(0.13)	(1.12)
Small-Cap Core	3,605	0.58	0.22	0.21	1.85	187.3	6.5	0.88	0.86	0.97
					(0.76)	(520.2)	(5.3)	(0.17)	(0.11)	(0.95)
Small-Cap Growth	4,258	0.62	0.19	0.19	2.01	161.5	6.7	0.92	0.84	1.27
					(0.71)	(516.5)	(5.3)	(0.07)	(0.09)	(0.84)
Small-Cap Value	3,330	0.58	0.25	0.18	1.87	185.1	6.0	0.92	0.80	0.66
					(0.73)	(608.6)	(4.6)	(0.08)	(0.14)	(0.56)
No Vanguard Entry during 1970-2005										
Mid-Cap Growth	4,507	0.63	0.20	0.17	1.95	231.4	8.5	0.86	0.81	1.31
					(0.75)	(741.2)	(9.2)	(0.10)	(0.11)	(1.03)
Mid-Cap Value	2,583	0.60	0.23	0.17	1.82	280.7	7.5	0.89	0.79	0.87
					(0.72)	(1,546.3)	(7.3)	(0.08)	(0.14)	(0.76)
Multi-Cap Growth	5,366	0.65	0.23	0.11	1.93	669.2	9.3	0.81	0.80	1.31
					(0.78)	(3,321.9)	(10.3)	(0.12)	(0.13)	(1.36)
Multi-Cap Value	6,678	0.57	0.29	0.14	1.72	401.2	9.5	0.78	0.84	0.69
					(0.74)	(1,764.1)	(11.6)	(0.13)	(0.12)	(0.64)

Table 5: Effects on Fees DID Estimate

This table reports the coefficients of the difference-in-differences regression with the basic specification in equation (5.2). Columns (1)-(5) use data at the fund share class level. The sample is limited to actively managed mutual funds in the investment styles that experienced Vanguard entry during 1970-2005. The dependent variable is the total fee calculated as Expense Ratio + (Front Load + Back Load)/7 and expressed in percentage points. Column (1) is the basic specification. Column (2) excludes institutional fund shares. Column (3) drops all observations beyond the fifth year post treatment. Column (4) is limited to the funds existing before VG entry, Column (5) keeps a balanced sample of funds with observations covering at least three years before entry and at least five years after entry. Columns (6) and (7) use data aggregated at the fund-channel level, which combines data on multiple share classes of the same fund in the same channel and same year. The dependent variable in Column (6) is the equal-weighted total fee for a fund-channel, and the dependent variable in Column (7) is the value-weighted average. Market Adjusted Return is the net raw return minus the market return. Year, category (investment style by investor type), channel, and fund fixed effects are controlled for in all regressions. Columns (6) and (7) also control for fund-channel fixed effects. All observations more than five years before Vanguard entry are dummied out. The regressions also allow for separate linear time trends by category. Standard errors clustered by fund family are in parentheses.

	(1)	(2)	(3) Total Fee %	(4)	(5)	(6) EW	(7) VW
Post VG 0-4 Yr	0.105***	0.096***	0.101**	0.091***	0.087**	0.091**	0.075**
	(0.034)	(0.034)	(0.042)	(0.035)	(0.036)	(0.040)	(0.034)
Post VG 0-4 Yr * Direct	-0.179***	-0.170***	-0.169***	-0.180***	-0.242***	-0.200***	-0.138***
	(0.044)	(0.044)	(0.043)	(0.044)	(0.045)	(0.044)	(0.044)
Post VG 0-4 Yr * Institutional	-0.128**		-0.014	-0.097**	-0.085*	-0.128***	-0.105**
	(0.050)		(0.046)	(0.047)	(0.047)	(0.048)	(0.042)
Post VG 5+ Yr	0.117***	0.111***		0.105**	0.075*	0.103**	0.085**
	(0.041)	(0.042)		(0.042)	(0.045)	(0.046)	(0.042)
Post VG 5+ Yr * Direct	-0.224***	-0.221***		-0.253***	-0.273***	-0.262***	-0.180***
	(0.048)	(0.048)		(0.045)	(0.049)	(0.045)	(0.047)
Post VG 5+ Yr * Institutional	-0.101**			-0.061	-0.008	-0.114**	-0.091**
	(0.050)			(0.049)	(0.054)	(0.048)	(0.045)
Pre VG 1 Yr	-0.008	-0.012	-0.003	-0.017	-0.030	-0.024	-0.009
	(0.037)	(0.037)	(0.040)	(0.036)	(0.037)	(0.031)	(0.036)
Pre VG 1 Yr * Direct	-0.051	-0.040	-0.063	-0.053	-0.026	-0.034	-0.033
	(0.045)	(0.045)	(0.043)	(0.045)	(0.044)	(0.038)	(0.041)
Pre VG 1 Yr * Institutional	0.003		0.054	0.021	0.042	0.017	0.006
	(0.043)		(0.042)	(0.041)	(0.044)	(0.036)	(0.039)
ln (Share Class Size) t-1	0.002	0.001	-0.013*	0.002	0.004		
	(0.006)	(0.008)	(0.007)	(0.007)	(0.008)		
ln (Fund Channel Size) t-1						-0.003	-0.001
						(0.007)	(0.007)
ln (Fund Size) t-1	-0.047***	-0.049***	-0.036***	-0.069***	-0.065***	-0.048***	-0.048***
	(0.009)	(0.012)	(0.011)	(0.011)	(0.012)	(0.009)	(0.009)
Age t	-0.007***	-0.008***	-0.006***	-0.007***	-0.007***	-0.008***	-0.012***
	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)
Market Adjusted Return t-1	-0.098***	-0.125***	-0.063**	-0.083***	-0.086***	-0.058***	-0.058***
	(0.020)	(0.023)	(0.026)	(0.024)	(0.024)	(0.022)	(0.022)
Turnover Ratio t-1	0.028***	0.027***	0.033***	0.034***	0.034***	0.031***	0.031***
	(0.006)	(0.006)	(0.010)	(0.008)	(0.009)	(0.006)	(0.006)
ln (Fund Family Size) t-1	-0.016**	-0.015*	-0.030***	-0.020*	-0.022**	-0.020***	-0.019***
	(0.007)	(0.008)	(0.011)	(0.011)	(0.010)	(0.006)	(0.006)
Year FE	Y	Y	Y	Y	Y	Y	Y
Category FE	Y	Y	Y	Y	Y	Y	Y
Channel FE	Y	Y	Y	Y	Y	Y	Y
Fund FE	Y	Y	Y	Y	Y	Y	Y
Fund-Channel FE	N	N	N	N	N	Y	Y
Linear Time Trends by Category	Y	Y	Y	Y	Y	Y	Y
SE Clustered by Fund Family	Y	Y	Y	Y	Y	Y	Y
Observations	26,982	22,420	7,972	13,393	9,419	16,805	16,805
R-squared	0.832	0.809	0.869	0.836	0.820	0.940	0.940

Table 6: Improved Specification

This table reports the coefficients of the difference-in-differences regression with the improved specification in equation (5.3), using data at the fund share class level. Sample is limited to actively managed mutual funds in the investment styles that experienced Vanguard entry during 1970-2005. The dependent variable is the total fee calculated as Expense Ratio + (Front Load + Back Load)/7 and expressed in percentage points. Column (1) includes the interactions of the share class size and the fund size with the pre, post dummies. Column (2) includes the interactions of all covariates with the pre, post dummies. The bottom panel shows the net treatment effects by channel, calculated from the coefficient estimates. When interactions with the controls are included, the effect on a channel is evaluated at the means of the covariates for this channel. All observations five years before Vanguard entry and earlier are dummied out. Standard errors clustered by fund family are in parentheses.

	(1)	(2)
	Total Fee %	
Post VG 0-4 Yr	0.074	0.083
	(0.070)	(0.094)
Post VG 0-4 Yr * Direct	-0.168***	-0.167***
	(0.041)	(0.041)
Post VG 0-4 Yr * Institutional	-0.126***	-0.131***
	(0.042)	(0.046)
Post VG 5+ Yr	0.132*	0.120
	(0.075)	(0.100)
Post VG 5+ Yr * Direct	-0.224***	-0.223***
	(0.045)	(0.045)
Post VG 5+ Yr * Institutional	-0.091**	-0.099**
	(0.043)	(0.049)
Year FE	Y	Y
Category FE	Y	Y
Channel FE	Y	Y
Fund FE	Y	Y
Linear Time Trends by Category	Y	Y
SE Clustered by Fund Family	Y	Y
ln Share Class or Fund Size t-1 * Pre, Post Dummies	Y	Y
All Control Variables * Pre, Post Dummies	N	Y
Observations	26,718	26,718
R-squared	0.834	0.834
Post VG 0-4 Yr, Retail Broker Channel	0.075*	0.077*
	(0.042)	(0.045)
Post VG 0-4 Yr, Retail Direct Channel	-0.073*	-0.073*
	(0.038)	(0.039)
Post VG 0-4 Yr, Institutional	-0.043	-0.046
	(0.030)	(0.031)

Table 7: Dynamics of Fee Responses

This table reports the coefficients of the difference-in-differences regression augmented with lead and lag year indicators relative to Vanguard entry. The dependent variable is the total fee calculated as Expense Ratio + (Front Load + Back Load)/7 and expressed in percentage points. Data are aggregated at the fund-channel level, which combines multiple share classes of the same fund-channel-year. The dependent variable in Column (1) is the equal-weighted total fee and in Column (2) is the value-weighted average. Covariates are the same as in Columns (6)-(7) of Table 5 and their coefficients are omitted. Year, category, and fund-channel fixed effects are controlled for in all regressions. All observations five years before Vanguard entry and earlier are dummied out. The regressions also allow for separate linear time trends by category. Standard errors clustered by fund family are in parentheses.

VARIABLES	(1) Total Fee EW	(2) Total Fee VW
Entry Year (-2)	-0.034	-0.039
	(0.031)	(0.031)
Entry Year (-1)	-0.031	-0.018
	(0.035)	(0.040)
Entry Year (0)	0.047	0.058
	(0.042)	(0.042)
Entry Year (1)	0.050	0.045
	(0.047)	(0.038)
Entry Year (2)	0.091*	0.066*
	(0.046)	(0.038)
Entry Year (3)	0.087*	0.060
	(0.048)	(0.042)
Entry Year (4)	0.093*	0.059
	(0.049)	(0.045)
Entry Year (5+)	0.095*	0.069
	(0.050)	(0.045)
Entry Year (-2) * Direct	0.011	0.027
	(0.040)	(0.039)
Entry Year (-1) * Direct	-0.031	-0.025
	(0.040)	(0.043)
Entry Year (0) * Direct	-0.137***	-0.115**
	(0.045)	(0.046)
Entry Year (1) * Direct	-0.140***	-0.097**
	(0.048)	(0.046)
Entry Year (2) * Direct	-0.198***	-0.129***
	(0.049)	(0.047)
Entry Year (3) * Direct	-0.234***	-0.151***
	(0.051)	(0.049)
Entry Year (4) * Direct	-0.239***	-0.150***
	(0.049)	(0.049)
Entry Year (5+) * Direct	-0.274***	-0.183***
	(0.045)	(0.047)
Entry Year (-2) * Institutional	-0.012	-0.007
	(0.046)	(0.045)
Entry Year (-1) * Institutional	-0.012	-0.024
	(0.053)	(0.055)
Entry Year (0) * Institutional	-0.095	-0.104
	(0.065)	(0.063)
Entry Year (1) * Institutional	-0.146**	-0.139**
	(0.073)	(0.067)
Entry Year (2) * Institutional	-0.201***	-0.176**
	(0.077)	(0.070)
Entry Year (3) * Institutional	-0.207**	-0.182**
	(0.082)	(0.077)
Entry Year (4) * Institutional	-0.213**	-0.183**
	(0.088)	(0.086)
Entry Year (5+) * Institutional	-0.203**	-0.184**
	(0.089)	(0.085)
Year FE	Y	Y
Category FE	Y	Y
Channel FE	Y	Y
Fund-Channel FE	Y	Y
Linear Time Trends by Category	Y	Y
SE Clustered by Fund Family	Y	Y

Table 8: Effects on Fee Components

This table reports the results from the difference-in-differences regression using the components of fees as the dependent variables. The sample is limited to actively managed mutual funds in the investment styles that experienced Vanguard entry during 1970-2005, and the data are at the share class level. Column (1) uses total fee as the dependent variable, calculated as Expense Ratio + (Front Load + Back Load)/7 and expressed in percentage points. Column (2) uses the marketing fee as the dependent variable, calculated as the amortized loads plus the annual distribution fee 12b-1 fee. Column (3) uses the non-marketing fee, which is the annual expense ratio minus the 12b-1 fee. Covariates are the same as in Table 5 and their coefficients are omitted. Year, category, channel and fund fixed effects are controlled for in all regressions. All observations five years before Vanguard entry and earlier are dummied out. The regressions also allow for separate linear time trends by category. Standard errors clustered by fund family are in parentheses.

	(1) Total Fee %	(2) Marketing Fee	(3) Non-Marketing Fee
Post VG 0-4 Yr	0.106***	0.051**	0.055**
	(0.032)	(0.021)	(0.027)
Post VG 0-4 Yr * Direct	-0.163***	-0.061**	-0.101***
	(0.041)	(0.024)	(0.037)
Post VG 0-4 Yr * Institutional	-0.127***	-0.102***	-0.025
	(0.043)	(0.033)	(0.032)
Post VG 5+ Yr	0.118***	0.058**	0.061*
	(0.039)	(0.025)	(0.033)
Post VG 5+ Yr * Direct	-0.207***	-0.105***	-0.103***
	(0.045)	(0.033)	(0.038)
Post VG 5+ Yr * Institutional	-0.099**	-0.072**	-0.027
	(0.043)	(0.036)	(0.038)
Year FE	Y	Y	Y
Category FE	Y	Y	Y
Channel FE	Y	Y	Y
Fund FE	Y	Y	Y
Linear Time Trends by Category	Y	Y	Y
SE Clustered by Fund Family	Y	Y	Y
Observations	26,975	26,975	26,975
R-squared	0.833	0.823	0.820

Table 9: Robustness Checks

This table presents the results of the robustness checks. Column (1) clusters the standard errors by category (investment style-by-investor type). Column (2) does two-way clustering by both category and year. The dependent variable in Column (2) is the residual of the total fee after partialing out fund fixed effects. Column (3) does not control for micro-level covariates. Column (4) does not include linear trends across categories. Column (5) calculates the total fee using five years as the amortization period for the loads. Column (6) uses nine years as the amortization period. Column (7) uses the annual expense ratio as the dependent variable. Column (8) controls for category-by-year fixed effects. Column (9) controls for channel-by-year fixed effects and the sample is limited to 1985-2005. Year, category, distribution channel and fund fixed effects are controlled for in all regressions. All observations five years before Vanguard entry and earlier are dummied out. All regressions, except for Column (4), allow for separate linear time trends by category. For columns (3)-(9), standard errors clustered by fund family are in parentheses.

	(1) Total Fee %	(2) Total Fee % Residual	(3) Total Fee %	(4) Total Fee %	(5) Total Fee % Loads 5-Yr Amort	(6) Total Fee % Loads 9-Yr Amort	(7) Exp Ratio %	(8) Total Fee %	(9) Total Fee %
Post VG 0-4 Yr	0.108**	0.106**	0.105***	0.091***	0.090***	0.119***	0.155***		0.069*
	(0.050)	(0.044)	(0.034)	(0.028)	(0.035)	(0.032)	(0.033)		(0.037)
Post VG 0-4 Yr * Direct	-0.164***	-0.146***	-0.176***	-0.160***	-0.110**	-0.194***	-0.299***	-0.187***	-0.037
	(0.054)	(0.043)	(0.039)	(0.041)	(0.045)	(0.040)	(0.040)	(0.042)	(0.051)
Post VG 0-4 Yr * Institutional	-0.129**	-0.129***	-0.105**	-0.111***	-0.105**	-0.142***	-0.187***		-0.022
	(0.050)	(0.049)	(0.041)	(0.040)	(0.048)	(0.041)	(0.039)		(0.043)
Post VG 5+ Yr	0.122**	0.120***	0.088**	0.101***	0.096**	0.136***	0.185***		0.077*
	(0.055)	(0.043)	(0.043)	(0.032)	(0.041)	(0.038)	(0.041)		(0.044)
Post VG 5+ Yr * Direct	-0.209***	-0.197***	-0.231***	-0.205***	-0.134***	-0.251***	-0.397***	-0.232***	-0.013
	(0.067)	(0.054)	(0.043)	(0.045)	(0.051)	(0.043)	(0.040)	(0.047)	(0.065)
Post VG 5+ Yr * Institutional	-0.102*	-0.102*	-0.029	-0.082**	-0.060	-0.126***	-0.208***		-0.020
	(0.057)	(0.062)	(0.050)	(0.042)	(0.048)	(0.042)	(0.046)		(0.050)
ln (Share Class Size) t-1	0.002	0.001		0.001	0.014**	-0.005	-0.030***	0.001	0.001
	(0.006)	(0.005)		(0.006)	(0.007)	(0.006)	(0.005)	(0.006)	(0.006)
ln (Fund Size) t-1	-0.047***	-0.001		-0.048***	-0.059***	-0.040***	-0.016**	-0.048***	-0.034***
	(0.011)	(0.004)		(0.009)	(0.010)	(0.009)	(0.008)	(0.009)	(0.009)
Age t	-0.007***	-0.004***		-0.007***	-0.006***	-0.008***	-0.010***	-0.007***	-0.007***
	(0.001)	(0.001)		(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Market Adjusted Return t-1	-0.097***	-0.075***		-0.106***	-0.093***	-0.100***	-0.108***	-0.123***	-0.099***
	(0.021)	(0.023)		(0.020)	(0.020)	(0.019)	(0.019)	(0.028)	(0.019)
Turnover Ratio t-1	0.028***	0.010***		0.028***	0.029***	0.027***	0.024***	0.027***	0.020***
	(0.004)	(0.000)		(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
ln (Fund Family Size) t-1	-0.016***	-0.010***		-0.016**	-0.015**	-0.017***	-0.019***	-0.017**	-0.010
	(0.006)	(0.003)		(0.007)	(0.007)	(0.006)	(0.006)	(0.007)	(0.006)
Year FE	Y	Y	Y	Y	Y	Y	Y	-	-
Category FE	Y	Y	Y	Y	Y	Y	Y	-	Y
Channel FE	Y	Y	Y	Y	Y	Y	Y	Y	-
Channel-by-Year FE	N	N	N	N	N	N	N	N	Y
Category-by-Year FE	N	N	N	N	N	N	N	Y	N
Fund FE	Y	N	Y	Y	Y	Y	Y	Y	Y
Linear Time Trends by Category	Y	Y	Y	N	Y	Y	Y	-	Y
SE Clustered by Fund Family	N	N	Y	Y	Y	Y	Y	Y	Y
SE Clustered by Category	Y	-	N	N	N	N	N	N	N
SE Clustered by Category and Year	N	Y	N	N	N	N	N	N	N
Observations	26,982	26,982	35,044	26,982	26,982	26,982	26,982	26,982	25,636
R-squared	0.832	0.352	0.793	0.832	0.833	0.827	0.754	0.835	0.834

Table 10: Aggregate Effects

This table reports the estimated effects in the aggregate by using group-level data aggregated at investment style-channel-year level. The dependent variable in Column (1) is the equal-weighted average total fee, and in Columns (2) and (4) is the size-weighted average total fee. Column (3) is estimated using a two-step procedure, in which I first regress share-class-level fees on micro covariates and group dummy variables (which gives "estimated group effects"), and then use the estimated group effects as the dependent variable in a difference-in-differences regression at the group level. Columns (1)-(3) drop the groups with fewer than ten observations, while Column (4) keeps all group observations. Year and category-by-distribution channel fixed effects are controlled for in all regressions. All observations five years before Vanguard entry and earlier are dummied out. The regressions also allow for separate linear time trends by category. Standard errors clustered by category are in parentheses.

	(1)	(2)	(3)
	Equal-Weight	Size-Weight	Adjust for Micro
Post VG 0-4 Yr	0.178**	0.146**	0.167***
	(0.068)	(0.062)	(0.053)
Post VG 0-4 Yr * Direct	-0.245*	-0.196*	-0.241***
	(0.129)	(0.106)	(0.070)
Post VG 0-4 Yr * Institutional	-0.252***	-0.155**	-0.213***
	(0.075)	(0.065)	(0.068)
Post VG 5+ Yr	0.211**	0.147	0.171**
	(0.099)	(0.089)	(0.077)
Post VG 5+ Yr * Direct	-0.257	-0.181	-0.234**
	(0.170)	(0.145)	(0.109)
Post VG 5+ Yr * Institutional	-0.316***	-0.126	-0.220**
	(0.100)	(0.089)	(0.090)
Pre VG 1 Yr	0.043	-0.025	0.016
	(0.029)	(0.048)	(0.038)
Pre VG 1 Yr * Direct	-0.063	-0.032	-0.064
	(0.038)	(0.061)	(0.074)
Pre VG 1 Yr * Institutional	-0.075**	-0.010	-0.059
	(0.027)	(0.048)	(0.046)
Year FE	Y	Y	Y
Style-Channel FE	Y	Y	Y
Linear Time Trends by Category	Y	Y	Y
SE Clustered by Category	Y	Y	Y
Observations	472	472	447
R-Squared	0.971	0.963	0.946

Table 11: Effects of Fund Flows

This table shows the estimated effects of index fund entry on the fund flows to actively managed funds. The sample is limited to actively managed mutual funds offered before the Vanguard index fund entry year in the styles that have experienced Vanguard entry during the sample period. The dependent variable in Columns (1)-(4) is the net fund flow rate, which is the dollar fund flow normalized by lagged asset. Columns (1)-(2) use data at the fund-channel level, and Columns (3)-(4) use data at the fund level. Observations where the lagged asset size is less than five million dollars are deleted. Columns (1) and (3) further drop the observations where the absolute value of the net fund flow rate is larger than 1, and Columns (2) and (4) drop the observations where the absolute value is larger than 0.8. The dependent variable in Column (5) is the natural log of fund-channel size, and the dependent variable in Column (6) is the natural log of the fee revenue collected by fund-channel. Covariates in Columns (1)-(4) include lagged log size, current age, log lagged fund family size, the coefficients on which are omitted. The performance measure is the lagged excess return over the benchmark. Covariates in Columns (5)-(6) include lagged log size, current age, lagged turnover ratio, and lagged log fund family size. The performance measure is the lagged market-adjusted return, which is the net raw return minus the market return. All observations five years before Vanguard entry and earlier are dummied out. The regressions in Columns (5)-(6) also allow for separate linear time trends by category. Standard errors in Columns (1)-(4) are clustered by fund family-year and in Columns (5)-(6) by investment style.

	(1)	(2)	(3)	(4)	(5)	(6)
	Fund-Channel Flow %		Fund Flow %		ln (Size)	ln (Revenue)
Post VG 0-4 Yr	-0.048**	-0.031*	-0.059***	-0.047***	-0.055	-0.007
	(0.020)	(0.018)	(0.017)	(0.016)	(0.042)	(0.050)
Post VG 0-4 Yr * Direct	-0.010	-0.026			-0.036	-0.139*
	(0.025)	(0.023)			(0.044)	(0.067)
Post VG 0-4 Yr * Institutional	-0.050	-0.058**	-0.054	-0.042	0.051	0.005
	(0.031)	(0.029)	(0.038)	(0.036)	(0.082)	(0.084)
Post VG 5+ Yr	-0.020	-0.006	-0.004	0.001	-0.009	0.046
	(0.025)	(0.022)	(0.023)	(0.021)	(0.060)	(0.074)
Post VG 5+ Yr * Direct	0.007	-0.004			-0.048	-0.219***
	(0.024)	(0.022)			(0.042)	(0.066)
Post VG 5+ Yr * Institutional	-0.111***	-0.104***	-0.159***	-0.122***	-0.005	-0.034
	(0.035)	(0.032)	(0.046)	(0.043)	(0.115)	(0.122)
Pre VG 1 Yr	-0.001	0.002	-0.029	-0.029	-0.030	-0.030
	(0.035)	(0.032)	(0.031)	(0.028)	(0.122)	(0.107)
Pre VG 1 Yr * Direct	-0.012	-0.024			0.002	-0.035
	(0.041)	(0.039)			(0.058)	(0.066)
Pre VG 1 Yr * Institutional	-0.048	-0.056	-0.037	-0.010	0.166	0.180
	(0.049)	(0.046)	(0.055)	(0.054)	(0.123)	(0.117)
Performance t-1	0.464***	0.396***	0.453***	0.382***	0.676***	0.644***
	(0.036)	(0.032)	(0.040)	(0.035)	(0.160)	(0.153)
Year FE	Y	Y	Y	Y	Y	Y
Category FE	Y	Y	Y	Y	Y	Y
Fund FE	-	-	Y	Y	-	-
Fund-Channel FE	Y	Y	-	-	Y	Y
Linear Time Trends by Category	N	N	N	N	Y	Y
SE Clustered by Fund Family-Year	Y	Y	Y	Y	N	N
SE Clustered by Category	N	N	N	N	Y	Y
Limit to Funds Existing Before Index Fund Entry	Y	Y	Y	Y	Y	Y
Observations	7,661	7,492	6,112	5,999	8,437	8,437
R-squared	0.380	0.390	0.360	0.360	0.962	0.959

Table 12: Heterogeneous Effects

Data are at the fund-channel level. The dependent variable in Columns (1)-(2), (5)-(6) and (9)-(10) is the asset-weighted total fee for the fund-channels. The dependent variable in Columns (3)-(4), (7)-(8) and (11)-(12) is the fund-channel flow rate calculated as the dollar flow normalized by lagged asset size. Observations where the absolute value of the flow rate is larger than one are deleted. The sample is limited to actively managed mutual funds in the styles that experienced Vanguard entry during the sample period, and limited to the funds that existed before the Vanguard index fund entry. The sample for Columns (1) and (3) contains the funds in large fund families–those with above-median fund family size ex ante (in the year before Vanguard entry, and the median is defined within year), and the sample for Columns (2) and (4) contains fund families with below-median size ex ante. The sample for Columns (5) and (7) contains the funds with above-median active share ex ante (here the median is defined within investment style-year), and the sample for Columns (6) and (8) contains those with below-median active share ex ante. The sample for Columns (9) and (11) contains the funds with above-median excess performance over benchmark ex ante (here the median is defined within investment style-channel-year), and the sample for Columns (10) and (12) contains those with below-median excess performance ex ante.Year, category, channel and fund-channel fixed effects are controlled for in all regressions. All observations five years before Vanguard entry and earlier are dummied out. The regressions also allow for separate linear time trends by category. Standard errors clustered by fund family are in parentheses.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
	Total Fee %		Fund-Channel Flow %		Total Fee %		Fund-Channel Flow %		Total Fee %		Fund-Channel Flow %	
	Fund Family Size				Active Share				Performance over Benchmark			
	Large	Small	Large	Small	High	Low	High	Low	High	Low	High	Low
Post VG 0-4 Yr	0.050	0.061	-0.056*	-0.050	0.125***	0.080	-0.013	-0.106**	0.059	0.010	-0.080*	-0.039
	(0.036)	(0.053)	(0.030)	(0.036)	(0.046)	(0.075)	(0.042)	(0.047)	(0.044)	(0.033)	(0.041)	-0.035
Post VG 0-4 Yr * Direct	-0.154***	-0.095*	-0.066*	0.067***	-0.150***	-0.119**	-0.005	0.016	-0.153***	-0.072*	0.056	-0.050
	(0.036)	(0.053)	(0.034)	(0.023)	(0.047)	(0.053)	(0.033)	(0.040)	(0.050)	(0.042)	(0.042)	(0.033)
Post VG 0-4 Yr * Institutional	-0.077**	-0.051	-0.046	0.121**	-0.150***	-0.080	0.089	0.046	-0.052	-0.042	0.033	0.011
	(0.035)	(0.052)	(0.056)	(0.045)	(0.050)	(0.071)	(0.056)	(0.050)	(0.053)	(0.035)	(0.051)	(0.041)
Post VG 5+ Yr	0.072	0.015	-0.054**	0.098***	0.175***	0.078	0.080	-0.007	0.025	0.056	-0.021	0.046
	(0.054)	(0.072)	(0.022)	(0.035)	(0.059)	(0.092)	(0.050)	(0.049)	(0.056)	(0.055)	(0.029)	(0.034)
Post VG 5+ Yr * Direct	-0.287***	-0.082	-0.060*	-0.010	-0.222***	-0.142*	-0.055	0.004	-0.191***	-0.196***	0.018	-0.088***
	(0.053)	(0.077)	(0.035)	(0.026)	(0.063)	(0.075)	(0.033)	(0.047)	(0.064)	(0.059)	(0.030)	(0.030)
Post VG 5+ Yr * Institutional	-0.070	-0.009	-0.069	0.026	-0.180***	-0.079	0.061	-0.088	-0.021	-0.076	-0.014	-0.078
	(0.060)	(0.073)	(0.070)	(0.080)	(0.065)	(0.098)	(0.090)	(0.085)	(0.069)	(0.060)	(0.076)	(0.065)
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Category FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Fund-Channel FE	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
Linear Time Trends by Category	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y	Y
SE Clustered by Fund Family	Y	Y	N	N	Y	Y	N	N	Y	Y	N	N
SE Clustered by Year	N	N	Y	Y	N	N	Y	Y	N	N	Y	Y
Observations	2,864	2,548	2,715	2,410	1,893	2,197	1,772	2,096	2,518	2,611	2,366	2,484
R-squared	0.946	0.893	0.377	0.387	0.947	0.926	0.330	0.391	0.934	0.917	0.386	0.353

Table 13: Flow Sensitivity to Fees

This table reports estimates from panel regressions where the dependent variable is yearly fund flow normalized by lagged asset. Total fee is calculated as annual expense ratio plus amortized loads, in percentage points. Observations in Columns (1)-(5) are at fund share class-year level. Regressions control for different performance measures. Columns (1)-(2) and (5) use lagged net market adjusted return (MAR) as the performance measure. Column (3) uses lagged pre-fee MAR. Column (5) controls for both lagged net MAR and lagged net 4-factor alpha, calculated using net monthly returns over the 24 months ending in December of that year. Column (4) estimates flow sensitivity to the annual expense ratio rather than the total fee. The samples for Columns (2), (5) and (6) include active funds only, while for other columns include both active and index funds. Column (6) uses data at the fund-channel level and controls for fund-channel fixed effects. To minimize the effects of outliers, observations where the absolute values of the percentage flows are larger than 1 or where the lagged asset sizes are less than five million dollars are deleted. All regressions control for investment style by year by channel fixed effects. All regressions include the following control variables and their interactions with channel fixed effects: lagged log fund share class size, lagged log fund family size, and current share class age in years. Coefficients on the control variables are not reported. Standard errors clustered by investment style-year are in parentheses.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable				Flow % t		
Performance measure	Net MAR	Net MAR	Pre-Fee MAR	Net MAR	Net MAR & Net 4-Factor Alpha	Net MAR
Total Fee t-1	-0.044***	-0.035***	-0.049***		-0.033***	0.028
	(0.006)	(0.006)	(0.006)		(0.006)	(0.018)
Total Fee t-1 * Direct	-0.038***	-0.021*	-0.038***		-0.022**	-0.082**
	(0.010)	(0.011)	(0.010)		(0.011)	(0.033)
Total Fee t-1 * Institutional	0.014	0.013	0.015		0.007	-0.116**
	(0.013)	(0.012)	(0.013)		(0.012)	(0.052)
Expense Ratio t-1				-0.056***		
				(0.008)		
Expense Ratio t-1 * Direct				-0.026**		
				(0.012)		
Expense Ratio t-1 * Institutional				0.016		
				(0.018)		
Performance t-1	0.696***	0.607***	0.692***	0.697***	0.419***	0.598***
	(0.072)	(0.066)	(0.072)	(0.072)	(0.062)	(0.079)
Performance t-1 * Direct	-0.204*	-0.181*	-0.201*	-0.204*	-0.127	-0.078
	(0.112)	(0.101)	(0.111)	(0.111)	(0.095)	(0.115)
Performance t-1 * Institutional	-0.189	-0.178*	-0.186	-0.190	-0.179*	-0.255**
	(0.117)	(0.107)	(0.117)	(0.116)	(0.094)	(0.116)
Net 4-Factor Alpha t-1					0.766***	
					(0.063)	
Net 4-Factor Alpha t-1 * Direct					-0.064	
					(0.101)	
Net 4-Factor Alpha t-1 * Institutional					0.135	
					(0.112)	
Investment Style-Year-Channel FE	Y	Y	Y	Y	Y	Y
Active Funds Only	N	Y	N	N	Y	Y
Standard Errors Clustered by Style-Year	Y	Y	Y	Y	Y	Y
Fund-Channel FE	N	N	N	N	N	Y
Observations	35,749	32,921	35,749	35,749	32,431	14,867
R-squared	0.187	0.183	0.186	0.187	0.205	0.508

Table 14: Effects on Flow-Fee Sensitivity

The data are at the fund-channel level, and the dependent variable is yearly dollar flow normalized by lagged asset size. Total fee is calculated as annual expense ratio plus amortized loads, in percentage points. Columns (1)-(2), (3)-(4) and (5)-(6) use data on the broker, direct and institutional channels, respectively. Regressions control for market-adjusted returns as the performance measure. Observations where the absolute values of the percentage flows are larger than 0.8 or where the lagged asset sizes are less than five million dollars are deleted. All regressions control for investment style-by-year fixed effects, as well as year fixed effects, category fixed effects and the pre-, post- dummy variables interacted with lagged total fee. Columns (2), (4) and (6) also include year fixed effects, category fixed effects and the pre-, post- dummy variables interacted with lagged market-adjusted return. All regressions include the following control variables: lagged log fund share class size, lagged log fund family size, and current share class age in years. Coefficients on the control variables are not reported. Standard errors clustered by investment style-year are in parentheses.

	(1)	(2)	(3)	(4)	(5)	(6)
			Flow % t			
	Broker		Direct		Institutional	
Post VG 0-4 Yr * Total Fee t-1	0.057*	0.072**	-0.005	0.011	-0.014	0.034
	(0.034)	(0.033)	(0.048)	(0.051)	(0.103)	(0.098)
Post VG 5+ Yr * Total Fee t-1	0.081**	0.081**	0.009	0.016	-0.056	-0.001
	(0.038)	(0.037)	(0.051)	(0.056)	(0.108)	(0.104)
Post VG 0-4 Yr * MAR t-1		0.100		-0.105		-0.646
		(0.200)		(0.225)		(0.404)
Post VG 5+ Yr * MAR t-1		0.373		-0.166		-0.730
		(0.227)		(0.288)		(0.492)
Year FE * Total Fee t-1	Y	Y	Y	Y	Y	Y
Style FE * Total Fee t-1	Y	Y	Y	Y	Y	Y
Fund-Channel FE	Y	Y	Y	Y	Y	Y
Standard Errors Clustered by Style-Year	Y	Y	Y	Y	Y	Y
Investment Style-Year FE	Y	Y	Y	Y	Y	Y
Observations	6,681	6,681	4,581	4,581	3,165	3,165
R-squared	0.554	0.573	0.505	0.523	0.505	0.520

Table 15: Effects on Active Management

This table estimates panel difference-in-differences regressions for the effect of index fund entry on the level of active management for retail funds 1980-2005. Observations are at the fund level. The dependent variable in Columns (1), (4)-(7) is the active share measure in Cremers and Petajisto (2009), which measures how much an active fund's portfolio deviates from the composition of its self-stated benchmark. The dependent variable in Column (2) is the tracking error. The dependent variable in Column (3) is the R-squared over the benchmark, which is calculated by regressing the net monthly share class returns over the past 24 months on the monthly returns of the benchmark index. Column (6) is limited to the sample of ex ante closet indexing funds (funds whose ex ante active share is below the median in its style). Column (7) is limited to the ex ante true active funds (funds whose ex ante active share is above the median in its style). Control variables whose coefficients are omitted include lagged log fund asset size, current fund age, lagged log fund family size, and lagged fund turnover. Year, investment style, investor type and fund fixed effects are controlled for in all regressions. The regressions also allow for separate linear time trends by category. All observations five years before Vanguard entry and earlier are dummied out. Standard errors clustered by fund family are in parentheses.

	(1) Active Share	(2) Tracking Error	(3) R-Squared on BM	(4) Active Share	(5) Active Share	(6) Closet Index	(7) True Active
Post VG 0-4 Yr	0.031***	0.008*	-0.018	0.012	0.031	0.036***	0.016
	(0.010)	(0.005)	(0.011)	(0.008)	(0.021)	(0.013)	(0.011)
Post VG 0-4 Yr * Closet Index				0.025**	0.027**		
				(0.010)	(0.011)		
Post VG 5+ Yr	0.016	0.008	0.013	-0.025	-0.029	0.020	-0.011
	(0.013)	(0.006)	(0.015)	(0.015)	(0.038)	(0.021)	(0.016)
Post VG 5+ Yr * Closet Index				0.050***	0.050***		
				(0.018)	(0.017)		
Fund FE	Y	Y	Y	Y	Y	Y	Y
Linear Time Trends by Style	Y	Y	Y	Y	Y	Y	Y
SE Clustered by Fund Family	Y	Y	Y	Y	Y	Y	Y
ln Fund Size t-1 * Pre, Post Dummies	N	N	N	N	Y	N	N
Observations	6,818	6,643	8,675	2,697	2,798	1,464	1,233
R-squared	0.870	0.671	0.730	0.816	0.821	0.796	0.798

Table 16: Summary Statistics of Index Funds vs. ETFs

This table reports the summary statistics of traditional index funds using data at the share class level over the sample period 1990-2005. For comparison, the statistics on ETFs are also reported. Total Fee is calculated as the annual expense ratio plus one-seventh of the sum of the front load and the back load. Share Class Size is the Total Net Assets in millions. Years since Inception are calculated using the first offer dates of the share classes. Turnover is the the fraction of a fund's holdings that has changed over a year.

	Broker			Direct			Institutional			ETF		
	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD
Total Fee %	1.30	1.35	0.68	0.47	0.35	0.41	0.34	0.32	0.20	0.22	0.20	0.07
Front Load %	1.12	0	2.08	0	0	0	0	0	0	0	0	0
Back Load %	1.22	0.50	1.71	0	0	0	0.07	0	0.33	0	0	0
Expense Ratio %	0.96	0.81	0.55	0.47	0.35	0.41	0.32	0.30	0.20	0.22	0.20	0.07
12b-1 Fee %	0.58	0.50	0.35	0.09	0.02	0.10	0.07	0.03	0.09	0.02	0.01	0.02
Share Class Size $ Million	366.1	48.6	1,060.8	2,839.2	213.4	10,495.7	880.3	170.8	3,252.2	3,257.8	586.2	7,727.5
Years since Inception	5.6	5.0	3.4	6.5	5.0	4.6	6.3	5.0	3.8	4.0	4.0	2.2
Turnover	0.57	0.10	3.33	1.11	0.14	4.93	0.22	0.13	0.24	0.19	0.14	0.17
Observations	4,135			1,898			2,541			725		

Table 17: Effects of ETF Entry on Traditional Index Fund Fees

The sample under study contains traditional index funds with quarterly data during 1990Q1-2005Q4. The dependent variable is the total fees of share classes in percentages. Column (1) is the basic specification. Column (2) drops all Vanguard index funds. Column (3) drops all observations that are more than three years before or four years after the treatment. Column (4) uses the basic specification but includes the interactions of the share class size and fund size with the pre, post dummies. Column (5) includes the interactions of all covariates with the pre, post dummies. Column (6) reports the coefficients from the regression including indicators for event years and shows the dynamics in the effects. Category-specific linear trends and fund fixed effects are included in all regressions. Standard errors clustered by fund family are in parentheses.

	(1)	(2)	(3)	(4)	(5)		(6)
Post ETF	0.13*	0.12	0.15***	0.17***	0.16	Entry Year (-1)	2.03
	(0.072)	(0.074)	(0.050)	(0.059)	(0.104)		(5.917)
Post ETF * Direct	-0.18**	-0.17*	-0.18**	-0.19***	-0.18**	Entry Year (0)	5.98
	(0.081)	(0.091)	(0.068)	(0.071)	(0.073)		(7.103)
Post ETF * Institutional	-0.13	-0.11	-0.17**	-0.15**	-0.15*	Entry Year (1)	13.46
	(0.087)	(0.098)	(0.072)	(0.073)	(0.076)		(8.199)
Pre ETF 1 Yr	0.00	-0.00	0.05			Entry Year (2)	14.51*
	(0.052)	(0.052)	(0.048)				(8.674)
Pre ETF 1 Yr * Direct	0.02	0.03	-0.04			Entry Year (3)	15.29*
	(0.060)	(0.071)	(0.053)				(8.748)
Pre ETF 1 Yr * Institutional	0.06	0.07	-0.01			Entry Year (4+)	16.32**
	(0.058)	(0.062)	(0.057)				(8.020)
						Entry Year (-1) * Direct	-6.00
							(7.657)
						Entry Year (0) * Direct	-10.63
							(8.987)
						Entry Year (1) * Direct	-23.99**
							(9.585)
						Entry Year (2) * Direct	-21.97*
							(11.371)
						Entry Year (3) * Direct	-24.91**
							(11.384)
						Entry Year (4+) * Direct	-35.05***
							(10.254)
						Entry Year (-1) * Institutional	4.56
							(7.207)
						Entry Year (0) * Institutional	-2.82
							(9.489)
						Entry Year (1) * Institutional	-18.29
							(11.152)
						Entry Year (2) * Institutional	-20.93*
							(11.477)
						Entry Year (3) * Institutional	-22.14*
							(11.292)
						Entry Year (4+) * Institutional	-21.25*
							(11.099)
Quarter FE	Y	Y	Y	Y	Y		Y
Category FE	Y	Y	Y	Y	Y		Y
Fund FE	Y	Y	Y	Y	Y		Y
Channel FE	Y	Y	Y	Y	Y		Y
Linear Time Trends by Category	Y	Y	Y	Y	Y		Y
SE Clustered by Fund Family	Y	Y	Y	Y	Y		Y
Observations	9273	8334	2386	9273	9273		9624
R-Squared	0.859	0.845	0.891	0.859	0.859		0.853

Figure 1: Growth of Index Funds and ETFs

This graph plots the growth of assets in index funds and ETFs. Data are from the CRSP Survivor-Bias-Free Mutual Fund Database. The solid line represents the market share of index funds in the open-end equity mutual fund (active fund + index fund) market. The dashed line represents the size of assets in the ETFs normalized by the size of the open-end equity mutual fund market. Data on ETF asset size are not available until 1998; thus, the relative ETF size is approximated with linear extrapolation from 1993-1998 (dotted line).

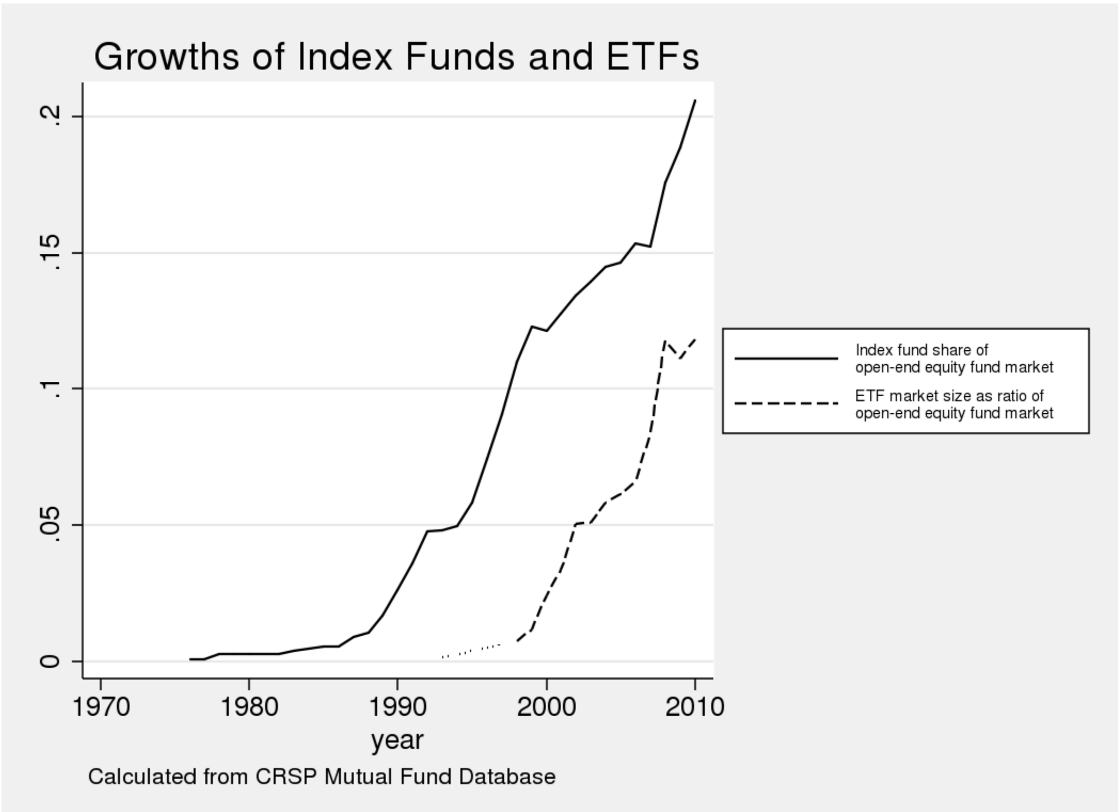


Figure 2: Equilibrium before Index Fund Entry

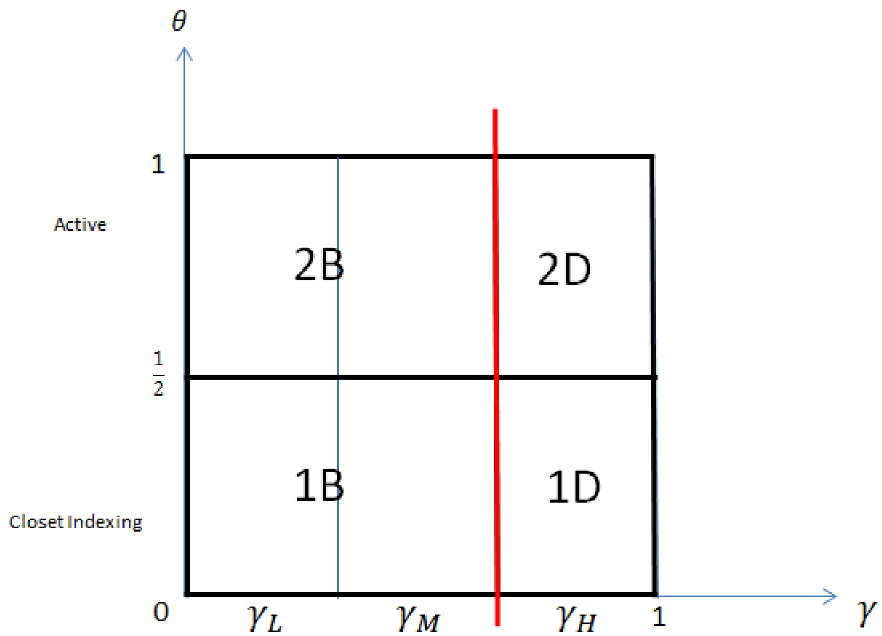


Figure 3: Equilibrium after Index Fund Entry

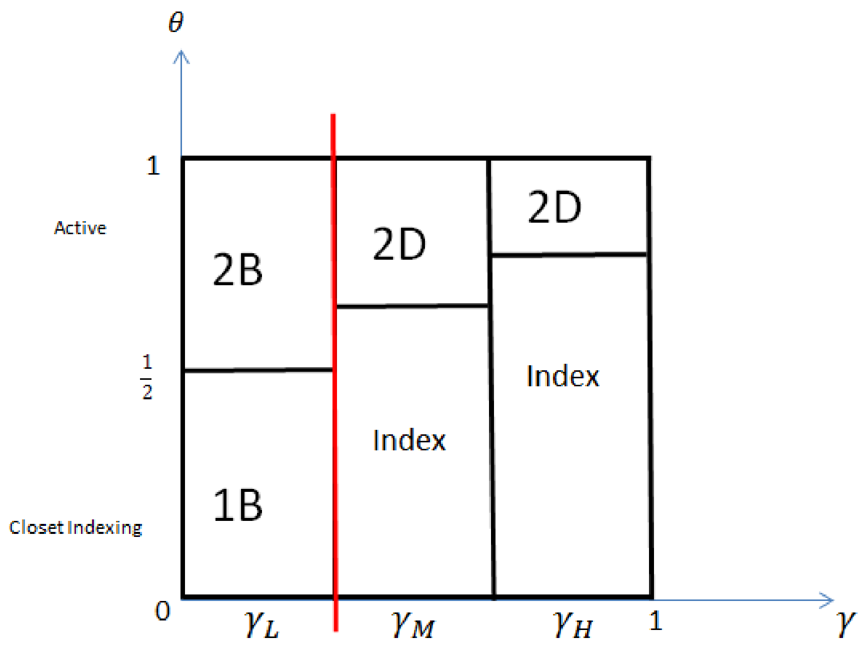


Figure 4: Equilibrium after ETF Entry

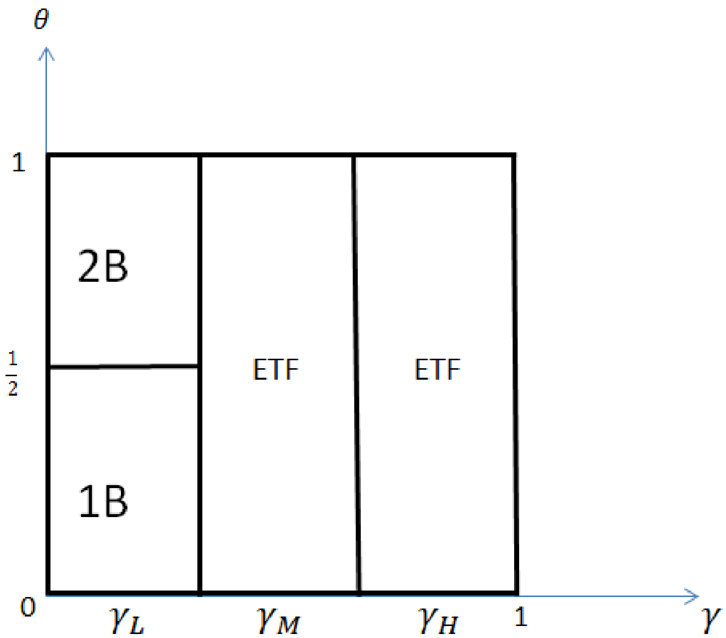


Figure 5: Dynamic Effect on Fees

The treatment effect dynamics are calculated from the coefficient estimates from the basic specification augmented with lead and lag year indicators relative to Vanguard entry. Vanguard enters between year -1 and year 0. The dependent variable is the total fee calculated as Expense Ratio + (Front Load + Back Load)/7 and expressed in percentage points. Data are at the fund-channel level, which combines multiple share classes of the same fund-channel-year. Year, category, and fund-channel fixed effects are controlled for in the regression. All observations five years before Vanguard entry and earlier are dummied out. The regressions also allow for separate linear time trends by category.

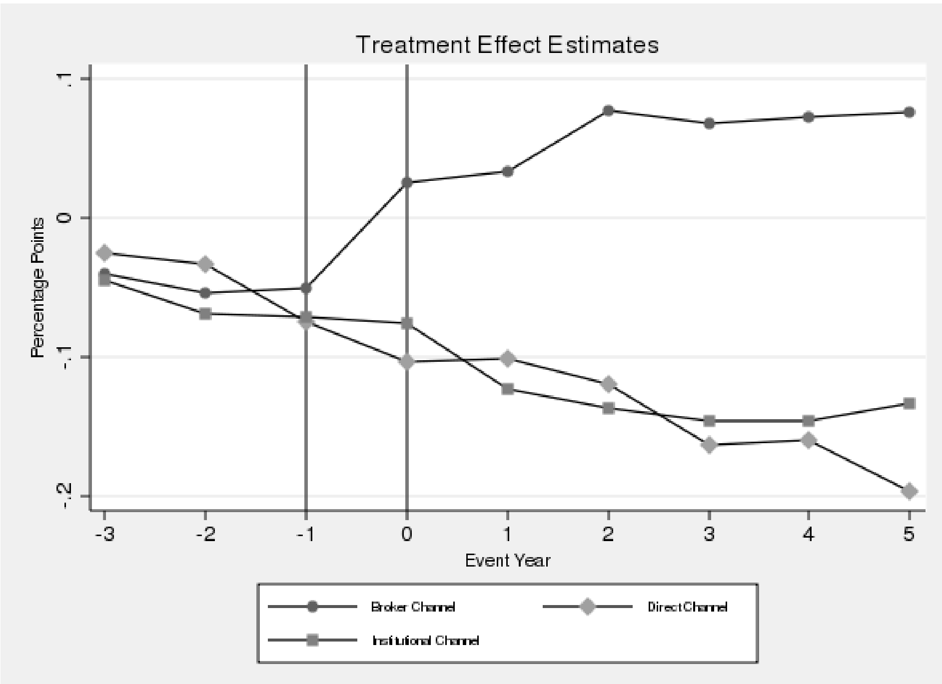
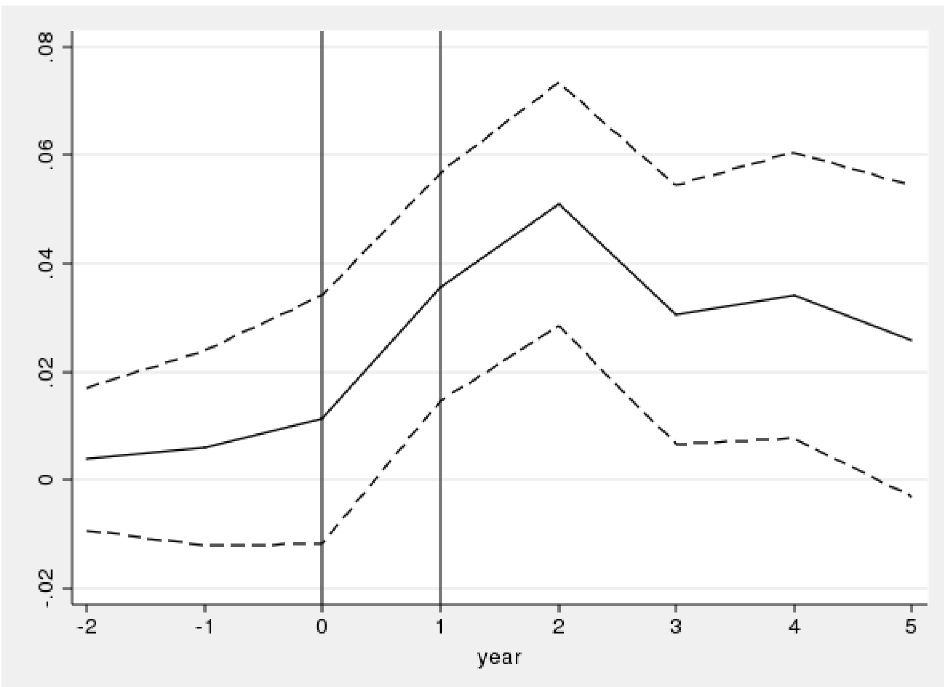


Figure 6: Dynamic Effect on Active Share

The dependent variable is the active share measure in Cremers and Petajisto (2009), which measures how an active fund's portfolio deviates from the composition of its self-stated benchmark. The treatment effect dynamics are the coefficient estimates from the DID regression augmented with lead and lag year indicators relative to Vanguard entry. The sample contains actively managed funds in the investment styles that experienced Vanguard entry during the sample period 1980-2005. Vanguard enters between year 0 and year 1. Data used are at the fund level. Control variables include lagged log fund asset size, current fund age, lagged log fund family size, and lagged fund turnover. Year, investment style, investor type and fund fixed effects are controlled for in all regressions. The regressions also allow for separate linear time trends by category. The dashed line shows the 95% confidence interval. Standard errors are clustered at the fund family level.



Appendix

Appendix 1 Proof of Proposition 1

An investor chooses the broker channel over the direct channel iff $\gamma \leq \frac{w}{p_B^* - p_D^*}$ (IC constraint). Market segmentation in the equilibrium according to Figure 2 requires that

$$\gamma_M(\frac{1}{\bar{\gamma}} - \frac{1}{\gamma_H}) < w < \gamma_H(\frac{1}{\bar{\gamma}} - \frac{1}{\gamma_H}). \tag{5}$$

The left inequality says that the M type needs to be insensitive enough so that they are willing to pay the price increment to have the fund-advice bundle. The right inequality requires the H type be sensitive enough; thus, they prefer the direct channel.

The non-profitable deviation conditions are the following.

In this equilibrium, each fund earns profit $\pi_1^* = \pi_2^* = \frac{1}{\bar{\gamma}} + \frac{1}{2\gamma_H}$. Because the solutions satisfy the first-order conditions for a quadratic problem, local deviations cannot increase the funds' profits. Ruling out non-local deviations–i.e., selling different products to the types holding the other fund's strategy constant–requires the additional restriction

$$w < 2\sqrt{\frac{(2\gamma_H + \bar{\gamma})(\gamma_H + 2\bar{\gamma})}{\bar{\gamma}\gamma_H}} - 6. \tag{6}$$

Whenever a higher type chooses the broker channel, all lower types also prefer the broker channel. Moreover, there does not exist a profitable deviation where one fund sells only 1D to all investors: At any price p_{1D} it would like to do so, there is a more profitable deviation where it sells 1B at $p_{1D} + \frac{w}{\gamma_H} - \epsilon$ to the same investors. Therefore, we need only consider the following two cases for potential profitable deviations.

a) If Fund 1 sells the fund-advice bundle 1B to all types, it has to lower price to compete with 2D for the H type. When 6 holds, this is not a profitable deviation for Fund 1 because, in order to sell to the H type who has low dollar valuation for financial advice, Fund 1 has to lose too much profit on the infra-marginal types M and L.

b) If Fund 1 sells 1D to the H and M types and 1B to the L type, it charges higher prices in both channels, mainly because the average investors in both channels are now less price-sensitive. When

$$w < \frac{\gamma_L + 3\gamma_M}{2\bar{\gamma}} + 2\sqrt{\frac{(2\bar{\gamma} + 4\gamma_H)\tilde{\gamma}}{\bar{\gamma}\gamma_H}}, \tag{7}$$

this is not a profitable deviation. Algebra shows that 6 implies 7. 7 also implies the second inequality in 5.

Appendix 2 Proof of Proposition 2

This appendix discusses the parameter restrictions under which the equilibrium in Proposition 2 holds. While this problem has a closed-form solution, given the number of parameters, to derive the full set of conditions for the solution is not tractable. Therefore, in the conditions below, I fix γ_H and γ_L and focus on examining the relationship between w and γ_M that allows the equilibrium.

In the equilibrium in Proposition 2, Fund 1's profit is

$$\Pi_1 = \frac{1}{2\gamma_L}.$$

Fund 2's profit is

$$\Pi_2 = \frac{c^2(1-\delta)^2\gamma_H^2\gamma_L + 9\gamma_M + \gamma_L(6 - c(1-\delta)\gamma_M)^2 + \gamma_H(9 - 12c(1-\delta)\gamma_L + 2c^2(1-\delta)^2\gamma_L\gamma_M)}{18\gamma_L(\gamma_H + \gamma_M)}.$$

a) If Fund 1 wants to keep the M type in the broker channel.

Fund 1 solves

$$\max_{p_{1B}}\left(\frac{1}{2} - \frac{\gamma_L}{2}p_{1B} + \frac{\gamma_L}{2}p_B^{*\prime} + \frac{1+w}{2} - \frac{\gamma_M}{2}p_{1B} + \frac{\gamma_M}{2}p_{2D}^{*\prime}\right)(p_{1B} - c)$$

s.t. $p_{1B} \leq \frac{w}{\gamma_M} + p_I$

Because the algebra for the no-deviation constraints becomes intractable, here, instead of using analytical expressions, I present numerical examples of the parameter restrictions under which the equilibria hold. Let $\gamma_H = 0.95$, $\gamma_L = 0.05$, $c = 6$ and $\delta = 0.3$. The set of (γ_M, w) that satisfies the condition in Proposition 1 – i.e., $\underline{w} < w < \overline{w}$, is given in Figure A1, Panel A. Including the additional no-deviation condition for Fund 1 in the new equilibrium, the set of parameters (γ_M, w) becomes the one illustrated in Panel B. The set of possible (γ_M, w) shrinks, excluding the area where γ_M, as a function of w, is low. Intuitively, this additional restriction means that given the value of financial advice, the M type is fee-sensitive enough, hence it is too costly for Fund 1 to keep the M type in the broker channel.

b) If Fund 2 wants to keep the M type in the broker channel.

Fund 2 solves

$$\max_{p_{2B};p_{2D}} \left(\frac{1}{2} - \frac{\gamma_L}{2} p_{2B} + \frac{\gamma_L}{2} p_B^{*'} + \frac{1+w}{2} - \frac{\gamma_M}{2} p_{2B} + \frac{\gamma_M}{2} p_I^{*'} + \frac{1}{2} - \frac{\gamma_H}{2} p_{2D} + \frac{\gamma_H}{2} p_I^{*'}\right)(p_{2B} - c)$$

s.t. $p_{2B} \leq \frac{w}{\gamma_M} + p_{2D}$.

Again let $\gamma_H = 0.95$, $\gamma_L = 0.05$, $c = 6$ and $\delta = 0.3$. Numerical calculations show that the no-deviation condition does not require additional restrictions on the parameters.

The set of parameters in Figure A1, Panel B also rules out all other deviations in the new equilibrium as stated in Proposition 2.

To explore the sensitivity of the range for (γ_M, w) that supports both equilibria to the values of γ_H and γ_L, Figure A2, Panel A shows the range of parameters (γ_M, w) for the equilibrium when $\gamma_H = 0.8$, $\gamma_L = 0.05$, $c = 6$ and $\delta = 0.3$. Figure A2, Panel B shows the range of parameters (γ_M, w) for the equilibrium when $\gamma_H = 0.95$, $\gamma_L = 0.1$, $c = 6$ and $\delta = 0.3$.

Appendix 3 Data Documentation

Sample Selection

The sample is constructed for years 1970-2005 from the CRSP Survivor-Bias-Free Mutual Fund Database and includes all U.S. domestic diversified equity mutual funds. The data are at mutual fund share class level. Sample selection builds on the procedure in Gil-Bazo and Ruiz-Verdu (2009) and includes the following steps. A mutual fund share class qualifies as a U.S. domestic diversified equity fund share class if it meets the following criteria in a given years. For 1962-1989, fund "policy" is common stock and the Wiesenberger Objective code is Growth, Growth-Income, or Maximum Capital Gains. For 1990-1991 (when there is no fund policy indicator), the Wiesenberger Objective code (which changed in 1990) is one of the following: Growth and Current Income, Long-term Growth, Maximum Capital Gains, or Small Capitalization Growth. For 1992-1997, the Strategic Insight objective code is one of Aggressive Growth, Mid-cap, Growth and Income, Growth, or Small-cap. From 1998 on, the Lipper objective code is one of Growth, Capital Appreciation, Growth and Income, Small-cap, Mid-cap, Micro-cap, or S&P 500 Index. I further dropped the observations before the official offer date and the share classes without reported initial offer dates.

Fund Identifiers

The fund identifiers are based on the crsp_cl_grp identifiers in CRSP, which contains data starting in 1998. The cl_grp identifiers are backfilled. About nine percent of the share classes do not have a CRSP fund identifier in any year. For those, fund identifiers were assigned using name parsing and a grouping method similar to that described in the CRSP documentation. Eventually, all share classes in the sample were assigned a fund identifier (those not grouped with others are assigned a fund identifier only for themselves).

Index funds and ETFs

In order to distinguish the index funds from the actively managed funds, I use the CRSP indicators, which start in 2008, combined with textual matching based on key words. In the textual matching method, funds with names containing the following strings are classified as index funds: Index, INDX, IDX, NASDAQ, BARRA, S&P, Dow Jones, Russell, ETF, exchange-traded, exchange traded. Funds with names containing the following strings are classified as ETFs: ETF, exchange-traded, exchange traded. All funds classified as index funds by their fund names but not marked in CRSP are all manually examined against their prospectus and classified accordingly. All funds marked as index-based or index-enhanced in CRSP (those with portfolios built upon indices but have made modifications) are not classified as index funds. All funds classified by their names but not marked in CRSP are all manually examined. Active ETFs, funds of funds, target date funds, or funds tracking unconventional indices (such as contrarian indices) are dropped. By the method above, I classified 287 index funds (553 share classes) and 74 ETFs in the basic sample.

Retail and Institutional Share Indicators

I construct an institutional indicator based on the fund share names. Share classes containing "Class I", "Class Y", "Class X", "Class K", "Institutional share", "Inst", "Trust Class", "Premier Class", "Fiduciary Class" or "Consultant Class" in their fund names are assigned to be institutional. [27]

Fund Family Identifiers

The fund family identifiers in this paper are based on the management codes in CRSP. The codes start in 1999. Between 1992 and 1999, management codes are filled based on the keywords in the management names. Before 1992, when no management names are available, the codes are backfilled. For the management names that do not have codes, I searched for key words of the names of fund families in the sample, such as "Fidelity", and manually assign the corresponding management codes. To cover the family names overlooked by the keyword search, I tabulated the lists of management names and codes and tried to fill in the codes whenever possible. For the fund families that cannot be matched to any codes following the above, I cleaned up the management names and assigned a separate code for each name. Management codes are then backfilled for each fund share class. In this way, I identified the management codes for 96 percent of all share classes. In some cases, different share classes of the same fund in the same year are assigned different management codes. This is usually due to fund mergers or acquisitions. If the share classes report the same fund name, I assigned the fund to the former fund family during the years that the old and new management codes overlap. If the share classes report different fund names, they belong

[27] As well as variations of the names, such as "Y Class" or "/Y" for "Class Y".

to different fund families and I leave the management codes unchanged.

Active Share and Tracking Error

I downloaded the active share data from Antti Petajisto's website http://www.petajisto.net/data.html, and merge them with the CRSP database using CRSP fundno. In the active share data set, the CRSP fundno is listed only for the largest share class of each fund. Therefore, I copied the data to other share classes within the same fund-year. The active share data start in 1980 and are available only for the funds that file the 13-F form with the SEC (those with assets > \$100 million). About 70 percent of the share classes observed in or after 1980 have available active share data.

Investment Categories

For the investment categories of funds, I used the Lipper classifications in the CRSP database based on portfolio holdings. I focused on 12 major categories of general U.S. domestic equity mutual funds: Large-Cap Core, Large-Cap Growth, Large-Cap Value, Mid-Cap Core, Mid-Cap Growth, Mid-Cap Value, Small-Cap Core, Small-Cap Growth, Small-Cap Value, Multi-Cap Core, Multi-Cap Growth, and Multi-Cap Value. Lipper has a separate category for S&P 500 index funds. For the purpose of my study, I treated this category as Large-Cap Core. Lipper classification is available from 1999. I used the earliest observed classification for each fund to backfill all earlier years of this fund's share classes. In this way, I obtained the category variable for 90 percent of the share classes in sample. I further filled the missing categories using the data on benchmark indices from Petajisto (2013), and the category variable became available for 92 percent of the share classes in sample. The funds whose categories could not be identified are not included in the sample.

Table A1 Hazard Model Estimate of Determinants for Vanguard Index Fund Entry

This table reports the estimates from a Cox proportional hazard model for the determinants of Vangaurd index fund entry into various categories of actively managed equity mutual funds. Columns (1)-(2) include all categories, while Columns (3)-(4) include only the categories that experienced Vanguard entry by 2005. All regressors are aggregated at the category level. Total fee, market-adjusted return (raw fund return minus the market return), share class size, turnover and return volatility are averages of all fund share classes in a category; in Columns (1) and (3), they are the equal-weighted averages, and in Columns (2) and (4), they are the asset-weighted averages (except for average share class size). Net flow % is the net flow percentage calculated for a category and is equal to $[CategorySize_t - CategorySize_{t-1} * (1 + WeightedAverageReturn_t)]/CategorySize_{t-1}$. Share of total market is a measure of relative market size for a category, calculated as the size of the category divided by the total asset size across all categories in a given year. Herfindahl is the Herfindahl–Hirschman Index for a category, calculated using market shares of the fund families. Institutional is an indicator that equals one if the category is for institutional investors.

	Hazard Model			
	EW	VW	EW	VW
Average Total Fee t-1	-0.030	-0.017	-0.024	-0.017
	(0.023)	(0.022)	(0.020)	(0.021)
Average Market Adjusted Return t-1	-0.636	-0.454	-0.904	-0.452
	(2.798)	(2.496)	(2.882)	(2.524)
Log (Average Share Class Size) t-1	-0.727	-0.718	-0.574	-0.676
	(0.619)	(0.669)	(0.596)	(0.658)
Average Share Class Age t-1	0.032	-0.185*	0.056	-0.145
	(0.138)	(0.107)	(0.132)	(0.105)
Average Turnover t-1	-2.228	-2.543	-1.756	-2.371
	(1.530)	(1.564)	(1.584)	(1.633)
Average Return Volatility t-1	0.085	0.059	0.047	0.031
	(0.187)	(0.196)	(0.185)	(0.194)
Net Flow % t-1	0.333	-0.069	0.294	-0.107
	(0.531)	(0.544)	(0.537)	(0.605)
Share of Total Market t-1	-0.889	18.769	-3.047	14.369
	(11.099)	(13.531)	(10.803)	(13.424)
Herfindahl t-1	-0.425	-0.186	-0.077	0.102
	(2.436)	(2.366)	(2.313)	(2.161)
Institutional	-2.585	-1.176	-2.075	-0.965
	(2.218)	(1.578)	(2.015)	(1.403)
All Markets	Y	Y	N	N
Only Markets Entered by VG by 2005	N	N	Y	Y
Observations	600	600	408	408

Table A2 Falsification Test

The dependent variable is the total fee calculated as Expense Ratio + (Front Load + Back Load)/7 and expressed in percentage points. The sample in Columns (1) is limited to the styles *not* entered by Vanguard by 2005, including Mid-Cap Growth, Mid-Cap Value, Multi-Cap Growth and Multi-Cap Value. In Column (1), the entry years for Mid-Cap Growth/Value are hypothetically assigned to be the same year of entry as Mid-Cap Core, and the entry years for Multi-Cap Growth/Value are assigned to be the same as that for Multi-Cap Core. Column (2) uses the sample of actively managed funds in the styles entered by Vanguard by 2005, and the hypothetical entry years are generated randomly. Covariates are the same as in Table 5 and their coefficients are omitted. Year, category, distribution channel and fund fixed effects are controlled for in all regressions. All observations five years before Vanguard entry and earlier are dummied out. The regressions also allow for separate linear time trends by category. Standard errors clustered by fund family are in parentheses.

	(1)	(2)
	Total Fee %	
Post VG 0-4 Yr	-0.020	0.022
	(0.063)	(0.032)
Post VG 0-4 Yr * Direct	-0.081	-0.042
	(0.057)	(0.049)
Post VG 0-4 Yr * Institutional	-0.066	-0.018
	(0.084)	(0.096)
Post VG 5+ Yr	0.032	-0.006
	(0.096)	(0.032)
Post VG 5+ Yr * Direct	-0.184***	-0.011
	(0.062)	(0.056)
Post VG 5+ Yr * Institutional	-0.150	-0.092
	(0.115)	(0.108)
Pre VG 1 Yr	-0.045	0.037
	(0.054)	(0.037)
Pre VG 1 Yr * Direct	0.089	-0.066
	(0.059)	(0.046)
Pre VG 1 Yr * Institutional	-0.016	0.060
	(0.065)	(0.097)
Year FE	Y	Y
Category FE	Y	Y
Channel FE	Y	Y
Fund FE	Y	Y
Linear Time Trends by Category	Y	Y
SE Clustered by Fund Family	Y	Y
Observations	13,403	26,982
R-squared	0.838	0.832

Table A3 Effects of ETF Entry on Components of Fees for Traditional Index Funds

The data contain traditional index funds with quarterly data during 1990Q1-2005Q4. The dependent variable in Column (1) is the total fee, calculated as the annual expense ratio plus amortized loads. The dependent variable in Column (2) is the expense ratio. The dependent variable in Column (3) is the marketing fee, defined as the amortized loads plus the annual 12b-1 fee. The dependent variable in Column (4) is the non-marketing fee, defined as the annual expense ratio minus the 12b-1 fee. Column (5) uses the annual 12b-1 fee as the dependent variable, and Column (6) uses the management fee (reported in CRSP). Quarter, category, channel, and fund fixed effects are controlled for in all regressions. Covariates, whose coefficients are omitted, include lagged log share class size, log fund size, log fund family size, current age, and lagged turnover ratio. Category-specific linear trends are included in all regressions. Standard errors clustered by fund family are in parentheses.

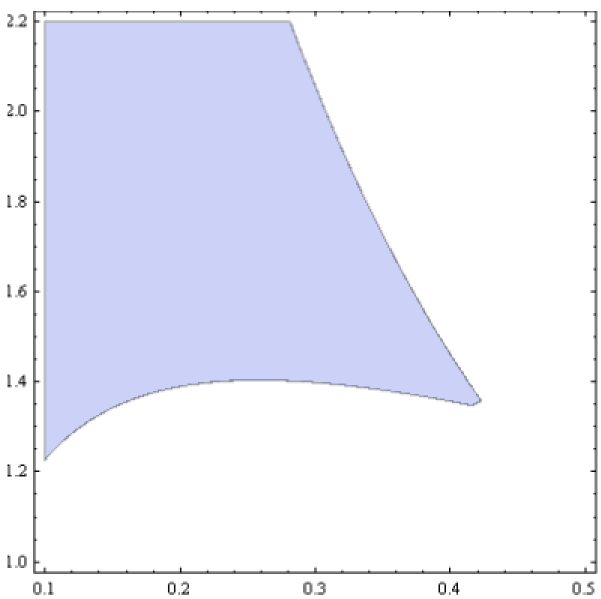
	(1) Total Fee	(2) Expense Ratio	(3) Marketing Fee	(4) Non-Marketing Fee	(5) 12b-1 Fee	(6) Management Fee
Post ETF	0.13**	0.05	0.13**	-0.00	0.05	-0.31
	(0.058)	(0.040)	(0.057)	(0.026)	(0.035)	(0.253)
Post ETF * Direct	-0.19***	-0.12**	-0.18***	-0.01	-0.11**	0.07
	(0.066)	(0.047)	(0.062)	(0.037)	(0.043)	(0.195)
Post ETF * Institutional	-0.15**	-0.02	-0.18***	0.03	-0.05	0.13
	(0.074)	(0.063)	(0.063)	(0.038)	(0.043)	(0.136)
Quarter FE	Y	Y	Y	Y	Y	Y
Category FE	Y	Y	Y	Y	Y	Y
Fund FE	Y	Y	Y	Y	Y	Y
Channel FE	Y	Y	Y	Y	Y	Y
Linear Time Trends by Category	Y	Y	Y	Y	Y	Y
SE Clustered by Fund Family	Y	Y	Y	Y	Y	Y
Observations	9273	9273	9273	9273	9273	7977
R-squared	0.859	0.806	0.810	0.835	0.675	0.497

Figure A1 Parameter Restrictions

$\gamma_H = 0.95$, $\gamma_L = 0.05$, $c = 6$ and $\delta = 0.3$

Panel A

Range of (γ_M,w) that satisfies the conditions in Proposition 1. γ_M on the horizontal axis, and w on the vertical axis.



Panel B

Range of (γ_M,w) that satisfies the conditions in Proposition 2.

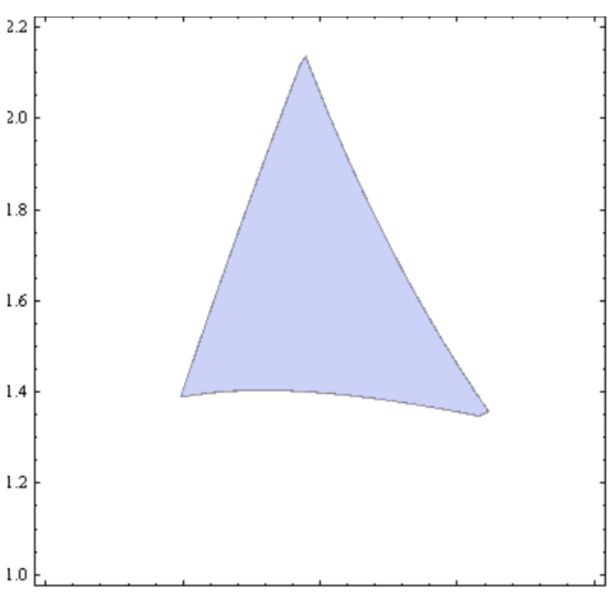
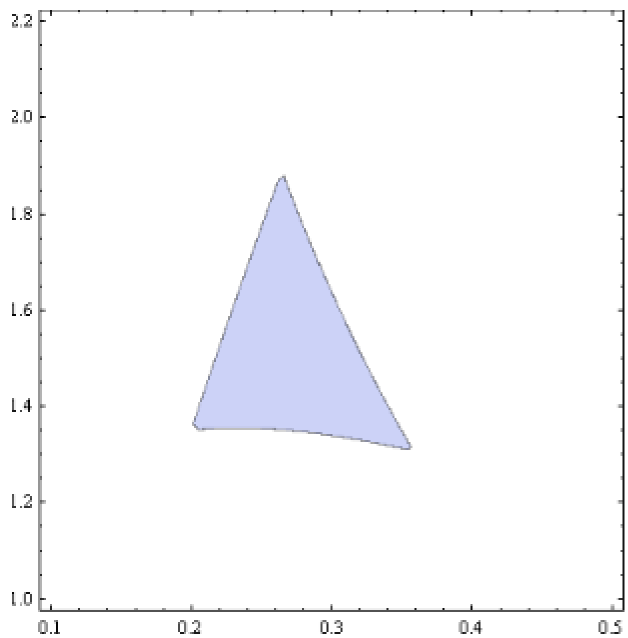


Figure A2 Parameter Restrictions for Proposition 2

Panel A

$\gamma_H = 0.8$, $\gamma_L = 0.05$, $c = 6$ and $\delta = 0.3$



Panel B

$\gamma_H = 0.8$, $\gamma_L = 0.1$, $c = 6$ and $\delta = 0.3$