RAND Journal of Economics
Vol. 42, No. 4, Winter 2011
pp. 776–804

Demand spillovers and market outcomes in the mutual fund industry

Alessandro Gavazza*

When consumers concentrate their purchases at a single firm, firms that offer more products than their rivals gain market share for all their products. These spillovers induce firms to offer a greater variety of products rather than lower prices, and a concentrated industry with few large firms can arise if spillovers are strong enough. This article presents a simple model that illustrates this mechanism explicitly. The empirical analysis documents strong demand spillovers in the retail segment of the U.S. mutual fund industry, in which fees are nontrivial, families offer many funds, and the market is quite concentrated. Instead, spillovers are weaker, fees are lower, families offer fewer funds, and the market structure is more fragmented in the institutional segment.

1. Introduction

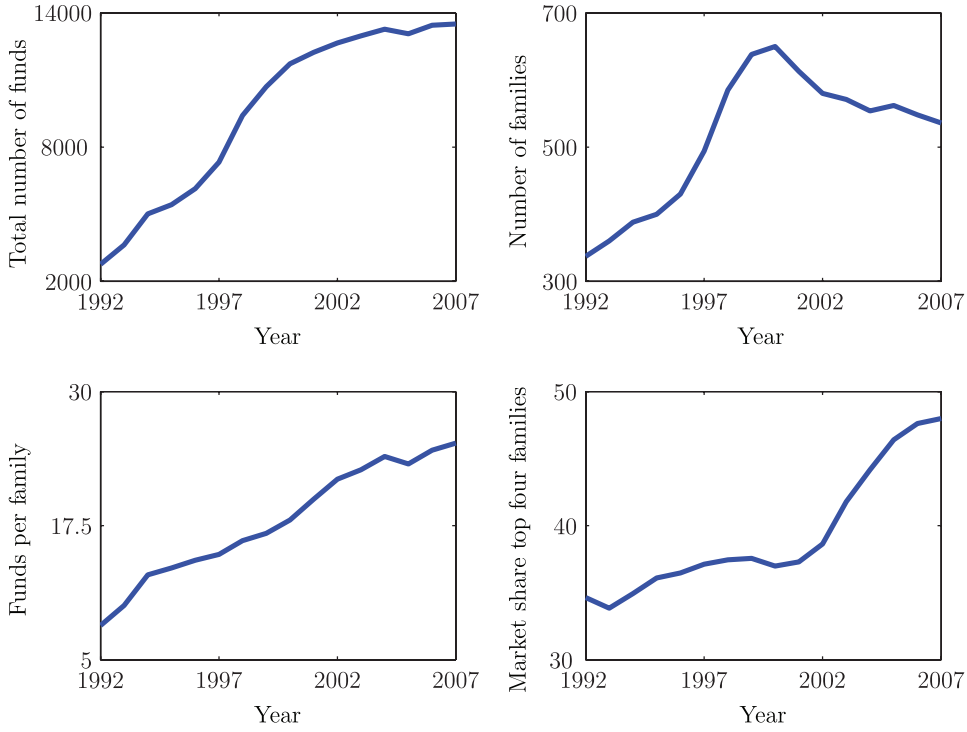
■ In many industries, buyers concentrate their purchases at a single supplier, because of transaction or switching costs. For example, many consumers shop at a single supermarket; several airlines acquire aircraft from a single manufacturer; and most individuals hold checking and savings accounts with the same bank. Thus, in such industries, firms' product lines are important strategic tools that vertically differentiate competing firms. The goal of this article is to investigate how these features—consumers deal with a single firm, and the number of products is a measure of vertical differentiation between firms—determine market conduct and market structure, with a special focus on the mutual fund industry.

With this goal in mind, the article sets up a model that builds on Sutton's (1991) seminal endogenous sunk costs theory. On the supply side, the key idea is that offering a large number of products affects fixed rather than variable costs. On the demand side, consumers prefer firms that offer a greater variety of products. In this setting, a firm that offers more products than its rivals gains market share for its other products as well (Klemperer and Padilla, 1997). These spillovers induce firms to compete by offering more and more products, in an escalation mechanism akin to the form of competition observed in markets characterized by network effects (Katz and Shapiro, 1985). The intensity of consumers' preferences for a firm's variety of products

*New York University; agavazza@stern.nyu.edu.
I am grateful to Heski Bar-Isaac, Allan Collard-Wexler, Michael Halling, Marcin Kacperczyk, Erik Kole, and seminar participants at the 2010 FIRS Conference, 2010 Professional Asset Management Conference at Erasmus University, University of Illinois at Urbana-Champaign, and University of Amsterdam for comments. The editor (Ali Hortaçsu) and two anonymous referees provided useful suggestions that substantially improved the article.

FIGURE 1

TRENDS IN THE MUTUAL FUND INDUSTRY, RETAIL FUNDS



critically determines the importance of spillovers and, thus, firms' conduct and market structure. If consumers' preferences for a firm's variety of products are strong, then a family gains a proportionally larger market share when offering more products than its rivals—that is, demand spillovers are strong. Hence, in equilibrium, firms offer a large number of products at relatively high prices. This mode of competition implies that incumbent firms incur high fixed costs, and new entrants must pay substantial setup costs to be able to compete with incumbents. Hence, the industry sustains only a few large firms and a natural form of market concentration arises. Instead, if consumers' preferences for a firm's variety of products are weak, then a firm gains a small market share when offering more products than its rivals—that is, demand spillovers are weak. Thus, each firm offers a more limited number of products, prices are lower, and the industry is more fragmented.

The mutual fund industry provides an ideal candidate for investigating the role of firms' variety of products and demand spillovers. First, industry data suggest that they play a role. Figure 1 documents some aggregate trends over the period 1992–2007 for the retail (i.e., non-institutional) segment of the industry. The growth of the industry in the last couple of decades has been explosive: the total number of retail funds available to investors grew from a little over 2,000 funds in 1992 to slightly under 14,000 in 2007, a sixfold growth in a 15 year period (top left panel). Concurrently, the total assets managed by retail funds grew from approximately 1.5 trillion dollars in 1992 to approximately 7.5 trillion dollars in 2007. Obviously, the industry's total growth attracted entry by some new families of funds.[1] On aggregate, the number of families

[1] A family of mutual funds is a group of funds marketed under a single brand name. Usually, the family is the distributor of its funds—that is, sells and redeems shares of its funds in transactions with investors—and an investment advisor as well. Examples of families are Vanguard, American Funds, and Fidelity. Examples of mutual funds are Vanguard Equity Income Fund, American Balanced Fund, and Fidelity Dividend Growth Fund.

increased by approximately 50% from 1992 to 2007 (top right panel), but the increase weakened throughout the period, in particular after 1999. Instead, the sixfold growth of the total number of funds was fueled by a fourfold increase in the average number of funds per family (bottom left panel). Moreover, the largest families grew disproportionately more than other families. As a result, the total market share of the largest four families almost doubled over the period 1992–2007 (bottom right panel).

Second, the mutual fund industry fits the key demand and supply assumptions of our model well.

(i) Most investors confine their mutual fund holdings to a single fund family, for two main reasons. First, employer-sponsored retirement plans frequently offer funds belonging to a single family. For example, Huberman and Jiang (2006), Elton, Gruber, and Blake (2006), and Cohen and Schmidt (2009) study samples of 401(k) plans and document that most of them offer funds of a single family. Second, "shopping" costs (i.e., transaction, search, and switching costs) induce investors to hold funds of a single family.[2] Some of these costs are monetary—a family often charges load fees when investors move assets out of the family, but not when they move assets within the family. Other costs could be psychological—the process of searching among a large number of funds across many families may be daunting.[3] As a result, Starks and Yates (2008) find that investors tend to buy funds from a single family even when they are investing through a discount brokerage firm.

(ii) The number of funds that a family offers plays a key role in determining choice among families. For example, when offering a retirement plan to a large number of employees with heterogeneous preferences over their portfolio choices, a family with a larger number of funds will better suit these employees. Similarly, shopping costs imply that moving money across funds of the same family is cheaper than moving money across families (Massa, 2003). The higher the number of funds in a family, the lower are investors' costs because more funds provide greater liquidity services. In summary, the number of funds is a key characteristic that vertically differentiates families.

(iii) The costs structure of the industry is well described by large fixed costs and low variable costs. Indeed, the main expenses to operate a fund—such as the compensation of the fund manager and administrative costs—do not grow proportionally with total net assets.

Third, the mutual fund industry is composed of two distinct segments—retail and institutional—that cater to different types of investors. In particular, institutional investors are usually larger and more sophisticated than retail investors (Section 3 provides more detail on the differences between segments). Hence, the difference between the benefits of finding a higher-performance or a lower-priced fund and the costs of switching across families is larger for institutional investors than for retail investors. Therefore, firms' variety of products and, thus, spillovers arguably play a different role in the two segments.

The empirical analysis starts by showing aggregate mutual fund data that reveal striking differences between the retail and institutional segments of the market. Specifically, in the retail segment, families offer more funds; funds have higher fees; the market is more concentrated; and the total number of funds is larger than in the institutional segment. These patterns are exactly the outcomes predicted by the theoretical model if preferences for firms' variety of products and spillovers are stronger in the retail segment than in the institutional segment. Thus, the empirical analysis seeks to measure and compare spillovers in the two segments of the market. In particular, using Center for Research in Security Prices (CRSP) data for the period 1999–2006, we examine whether families that offer a larger number of funds or a larger number of categories of funds

[2] Several authors document the importance of such costs in the mutual fund industry (e.g., Sirri and Tufano, 1998; Massa, 2003; Hortaçsu and Syverson, 2004). As the existence of these costs has already been established, this article does not offer direct evidence on them. Rather, the article derives implications of these costs for competition in the industry.

[3] Klemperer (1995) offers a rich illustration of these shopping costs.

gain a more-than-proportional market share. The main challenge to this analysis is to properly distinguish demand for firms' product portfolios from unobserved demand factors that may affect the relationship between the number of funds/categories and market share. Hence, as in a standard demand and supply estimation, in order to identify *demand* for firms' product varieties, we employ variables that shift the number of funds offered by a family for *supply* reasons. More precisely, an influential literature has empirically documented the spatial agglomeration of firms belonging to the same industry (for a summary, see Fujita, Krugman, and Venables, 1999), and a few articles have shown that firms located in these areas where an industry concentrates are, on average, larger than firms in the same industry outside such areas (for evidence across industries, see Holmes and Stevens, 2002; for evidence on the mutual fund industry, see Christoffersen and Sarkissian, 2009). Thus, everything else equal, the *supply* of funds/categories should be higher if a family is located where employment in the financial sector is higher. Hence, we retrieve from CRSP the zip code of the headquarters of each family, and we match it to the total employment in the financial sector (and to the total number of establishments) of the corresponding county, retrieved from the County Business Patterns, a U.S. Census database, to obtain a measure of the density of the financial sector relative to other industries.

We find clear evidence that demand for firms' variety of products and spillovers is stronger in the retail segment than in the institutional segment of the market. More precisely, a family that offers 10% more funds (categories) than its rivals has a 4.8% (10.1%) higher average per-fund market share in the retail segment, but does not gain a higher average per-fund market share in the institutional segment. We also perform our analysis at the fund and at the category level, and we find even stronger results. We further check the robustness of our results to potential concerns about the validity of the instruments using an alternative set of instrumental variables, and we find identical results.

The article makes a number of contributions. First, the mechanisms identified in this article are not unique to the mutual fund industry, but also help explain the role of firms' product portfolios and spillovers in a wide range of markets. In many industries, the largest firms are frequently the most successful in launching new products or entering new markets, and this article is among the first to investigate market outcomes when consumers value firms' product-line breadth. More generally, the article illustrates how markets operate when vertical differentiation between firms is important. Second, this article is one of the first to investigate market conduct and market structure in the mutual fund industry, an important industry that manages a large fraction of retirement savings in the United States. Although there have been other important trends in the mutual fund industry in recent years—such as the growth of index funds and exchange-traded funds (ETFs)—the article connects the industrial organization of the U.S. mutual fund industry with its largest component of demand—retirement accounts (Cohen and Schmidt, 2009). In doing so, the article provides a coherent explanation for why, given the industry's total growth, the entry of new families has been limited and the introduction of new funds by incumbent families has been more substantial. The model also provides a simple economic rationale for the limited price competition observed in the industry and for the existence of a large number of funds, both of which have been viewed as "puzzles" in the mutual fund literature (Gil-Bazo and Ruiz-Verdú, 2009; Massa, 1997). The economic mechanism proposed also matches well with key features of the industry and comments from market participants. For example, *The No-Load Fund Investor*,[4] a popular investment newsletter, writes: "The tremendous increase in the number of funds has been propelled by investor demand and by the need for fund groups to each offer a complete array of funds".

2. Related literature

■ This article contributes to several strands of the literature. First, an influential literature studies the determinants of market structure. Most theoretical explanations focus on supply-

[4] www.noloadfundinvestor.com.

side reasons. Sutton's seminal work highlights how endogenous sunk costs—such as advertising and research and development (R&D)—affect market structure and concentration (Sutton, 1991, 1998). Ellickson (2006) provides an interesting empirical application of Sutton's theory to the supermarket industry, in which competition in large markets induces supermarkets to expand the array of products offered to consumers, thereby reducing the entry of new firms. Berry and Waldfogel (2010) investigate the relationship among market size, product quality, and market concentration in two contrasting industries—daily newspapers and restaurants—that differ in whether quality is produced mainly with fixed or variable costs, respectively. The present article displays some key differences from these empirical articles: whereas they exploit variation in market size to understand the nature of supply (i.e., sunk costs), we focus on two segments of the same industry that differ in their demands—more precisely, in their valuations of the endogenous characteristic supplied through fixed costs.

Some articles use demand-side arguments to explain patterns of market concentration and market power. The most closely related articles are those that investigate the effects of switching costs on firms' incentives to offer multiple products (Klemperer, 1995; Klemperer and Padilla, 1997). Starting with Borenstein (1989, 1991), several empirical articles have investigated the "hub premium" in airline markets—whereby an airline with a dominant presence at an airport obtains a higher market share and charges higher prices on *all* trips originating from the dominant airport—and the role of switching costs, such as frequent-flyer programs, in generating this premium (Lederman, 2008). Also related is the literature on network externalities and, in particular, articles that investigate the effects of product compatibility on market outcomes (Katz and Shapiro, 1985, 1994; Matutes and Regibeau, 1988; Economides, 1989)—more precisely, the empirical articles that consider the effects of ATM surcharge fees on banks' optimal ATM network size and deposit account pricing (Knittel and Stango, 2008, 2009; Gowrisankaran and Krainer, 2011). One contribution of the present article is to empirically document that the mutual fund industry shares some key features with network industries.

This article contributes to a second strand of literature on multiproduct firms and product-line breadth. The theoretical literature suggests that firms offer multiple products to prevent the entry of rival firms (Schmalensee, 1978; Shaked and Sutton, 1990) and to increase their market share and profitability (Lancaster, 1979). A few articles empirically investigate the impact of a product-line extension, finding that firms are able to increase prices and sales once they extend their product line (Kadiyali, Vilcassim, and Chintagunta, 1998; Draganska and Jain, 2005). Unlike these articles, we illustrate how an industry's institutional arrangements can generate demand for firms' varieties, as well as their implications for competition, industry equilibrium, market structure, and concentration.

Finally, the article is related to the literature on mutual funds. Most articles analyze mutual funds' returns and/or portfolios, and only recently have a few articles focused on the role of families and on the industrial organization of the industry. Khorana and Servaes (1999) empirically examine the determinants of mutual fund starts and find that economies of scale and scope, the family's prior performance, and the overall level of funds invested are the main factors that induce families to set up new funds. Khorana and Servaes (2007) investigate the determinants of market shares and document that families that charge lower fees, perform better, and start more funds relative to the competition have a higher market share. Massa (1998, 2003) argues that families use market segmentation and fund proliferation to exploit investors' heterogeneity, limit competition, and increase market coverage. Wahal and Wang (2011) examine the impact of the entry of new funds on incumbents, and show that competition lowers management fees and investor flows. Nanda, Wang, and Zheng (2004) and Ivković (2004) investigate the importance of spillovers arising from the stellar performance of other funds in the same family. Hortaçsu and Syverson (2004) examine how search costs affect the proliferation of mutual funds and the dispersion of their fees. Pollet and Wilson (2008) investigate how funds and families respond to asset growth, finding that family growth is correlated with the introduction of new funds and that families with more funds diversify less rapidly as they grow. Park (2008) examines

how advertising affects the market structure of the mutual fund industry. Some of these articles mention demand-side spillovers (in particular, Massa, 2003), but none of them considers the supply side of the market, thereby characterizing industry equilibrium, market structure, and concentration.

3. Background: retail and institutional funds

■ The purpose of this section is to shed light on the key differences between the retail and institutional segments of the mutual fund industry. For a more thorough description of the industry, see Gremillion (2005).

Retail mutual funds are aimed toward individual investors—that is, households—and tax-advantaged individual retirement accounts have been instrumental in the growth of the retail segment in the last decades.[5] According to the Investment Company Institute (ICI, 2008) , many of today's mutual fund owners were introduced to mutual fund investing through retirement plans. For example, 59% of mutual fund-owning households indicated that they purchased their first fund through an employer-sponsored retirement plan, and that fraction increases to 68% for households that purchased mutual funds after 2000. Thus, retirement plans are, today, the most common source through which individuals invest in mutual funds: 51% of households that owned mutual funds viewed retirement plans as their main fund-purchase source (ICI, 2008). As a result, Cohen and Schmidt (2009) report that in 2004, more than 60% of new flows into non-money market mutual funds were due to retirement accounts.

Two main types of individual retirement plans exist: (i) employer-sponsored defined-contribution plans, such as 401(k) plans, and (ii) IRA/Keogh plans. Both types have an almost 50% share of total retirement assets, but they differ in several aspects.[6] An employer-sponsored plan is a benefit an employer provides to its employees. Thus, the employer initiates the plan, chooses the plan rules (eligibility, vesting, etc.), and arranges for its administration, commonly outsourcing it to an external organization, such as a mutual fund company, an insurance company, or an HR-benefit firm. The administrator is in charge of setting up an individual account for each plan participant, recording the participant's investment choices, buying and selling shares in particular mutual funds, and so forth. Although the function of administrator and of investment provider are sometimes separate, the same company more frequently provides both. Gremillion (2005) explains: "Administering a defined contribution plan, particularly performing participant record keeping, costs a great deal, often more than can be recovered in administration fees. Mutual fund companies that perform defined contribution record keeping—Fidelity, Vanguard, Putnam, etc.—do so primarily to gather assets into the funds, not to make money on record keeping." Indeed, Cohen and Schmidt (2009) find that inclusion in a 401(k) plan has a large positive effect on a family's inflows. The employer and the administrator negotiate over the investments available to the employees, thus restricting employees' choice set. For example, until recently, Vanguard allowed plans for which it performed record keeping to choose Vanguard mutual funds only. Indeed, Huberman and Jiang (2006), Elton, Gruber, and Blake (2006), and Cohen and Schmidt (2009), using different samples, document that most 401(k) plans offer funds of a single family only. Hence, a fund family with a larger number of funds is appealing to employers with a large number of employees, who may have heterogeneous preferences over their portfolio choices (because of age, income, or idiosyncratic taste). For instance, Agnew, Balduzzi, and Sundé n (2003) study allocations in retirement accounts and find that equity allocations are higher for

[5] Sialm and Starks (2009) investigate whether the characteristics, investment strategies, and performance of mutual funds held by taxable investors differ from those of funds held primarily in tax-deferred retirement accounts.

[6] Total individual-account retirement assets amounted to $6.767 trillion in 2004, according to the Employee Benefit Research Institute (2008). These assets were almost equally split between the two main forms of retirement plans: $3.384 trillion in employer-sponsored defined-contribution plans, such as 401(k) plans, and $3.383 trillion in IRA/Keogh accounts.

males, married investors, younger investors, investors with higher earnings, and those with more seniority on the job.[7]

An IRA plan, however, is not linked to employment, and any individual can open such a plan at a mutual fund company or a brokerage firm. Nonetheless, if the individual chooses to open an IRA with a mutual fund company, investment choices are still restricted exclusively to a family's funds. Similarly, if the individual chooses to open an IRA with a brokerage firm, families' offerings still affect investment choices through load fees.

Institutional investors—nonfinancial businesses, financial institutions, nonprofit organizations, state and local governments, and funds holding mutual fund shares—mainly purchase fund shares directly from fund companies that often create special share classes or funds expressly for them. Nonfinancial businesses and financial institutions use mutual funds primarily as a tool to manage their cash, thus investing the majority of their assets in money market funds. Instead, the investments of nonprofit organizations, state and local governments, and funds holding mutual fund shares are more evenly split across the main categories of funds (equities, fixed-income, hybrid, and money market funds). Institutional investors are larger and, on average, more sophisticated than individual investors (Hortaçsu and Syverson, 2004). Moreover, their investment choices are not restricted to a single family, and their costs of switching across families are small relative to the gains of finding a higher-performance or a lower-priced fund.

Overall, these considerations suggest that the number of funds that a family offers plays a different role in determining market shares in the retail and in the institutional segments. Thus, we expect that demand spillovers differ across the two segments. The next section presents a simple model that investigates how these demand spillovers affect market structure and market conduct. The model will guide the empirical analysis of Section 5.

4. A simple model

■ In this section, we introduce a simple model that adapts the theoretical framework of Sutton (1991) to the mutual fund industry. The main goal of the model is to investigate in the simplest way how the demand for a firm's product portfolio shapes competitive outcomes in the industry. More precisely, we make the simplest assumptions on the demand side of the model (assuming, in particular, that the decision maker chooses funds of a single family, rather than deriving it from first principles—i.e., shopping costs), and focus on two key families' supply choices: the number of funds offered and the fees. In turn, these choices determine the equilibrium number of families and, thus, market concentration. The main advantage of this setup is that it delivers testable implications that have intuitive empirical counterparts.

□ **Demand.** Consider a decision maker selecting a mutual fund family—for example, an employer with a large number of employees with heterogeneous preferences over portfolio choices. The decision maker chooses among families of funds that are vertically differentiated, differing by the number of funds offered N_j and the price p_j. The decision maker's indirect utility from choosing family j is equal to

$$u_j = \alpha_N \log N_j - \alpha_p p_j + \epsilon_j.$$

The parameter α_N measures the importance of demand spillovers: a higher α_N makes the number of funds N_J that a family offers more important for the decision maker. ϵ_j is a preference shock,

[7] Huberman and Jiang (2006) report these additional facts about investors' choice among retirement funds: "First, participants choose to invest their savings in a small number of funds—typically no more than three or four—regardless of the number of funds their plans offer. Second, a substantial fraction of participants tend to allocate their contributions evenly among the funds they choose. Third, there is little relation between the proportion of contributions that participants allocate to equity funds (equity allocation) and the proportion of equity funds that their plans offer (equity exposure)." See also Choi, Laibson, and Madrian (2004) for the impact of 401(k) plan design on savings outcomes.

assumed to be i.i.d. across consumers and families, distributed according to the type I extreme value distribution (McFadden, 1974). Thus, the market share of family j is equal to

$$s(N_j, p_j) = \frac{\exp(\alpha_N \log N_j - \alpha_p p_j)}{\sum_l \exp(\alpha_N \log N_l - \alpha_p p_l)}. \tag{1}$$

□ **Supply.** The industry is populated by a large number of homogeneous potential entrants (families) that compete by offering funds and by charging fees. A family j offering N_j funds has a cost function equal to

$$C(N_j) = K + FN_j, \tag{2}$$

where $K > 0$ is a setup cost and $F > 0$ is a fixed cost per fund. Thus, the cost function $C(N_j)$ assumes that fixed costs are strictly positive (so there are always economies of scale) and that the introduction of each fund affects families' fixed costs exclusively.

Using the market share (1) and the cost function (2), we obtain family profits as

$$\pi(N_j, p_j) = p_j s(N_j, p_j) M - C(N_j),$$

where M is the exogenous size of the market. Each family j chooses the number N_j^* of funds and the price p_j^* that maximize its profits $\pi(N_j, p_j)$. The first-order conditions for optimality are

$$p_j : Ms_j - p_j^* M \alpha_p s_j (1 - s_j) = 0$$
$$N_j : \frac{p_j^* M \alpha_N s_j (1 - s_j)}{N_j^*} - F = 0,$$

where $s_j = s(N_j^*, p_j^*)$.

Free entry determines the number S^* of active families in equilibrium. Thus, ignoring integer constraints, free entry drives down profits to zero—that is, $\pi(N_j^*, p_j^*) = 0$.

Proposition 1. In a symmetric equilibrium, in a large market $M \to +\infty$:

(i) if $\alpha_N > 1$, the number N^* of funds offered by each family is of the same order as the size of the market M, the price p^* is equal to $\frac{\alpha_N}{\alpha_p}$, and the number S^* of families is equal to $\frac{\alpha_N}{\alpha_N - 1}$;
(ii) if $\alpha_N \leq 1$, the number N^* of funds offered by each family is equal to $\frac{\alpha_N K}{(1-\alpha_N)F}$, the price p^* is equal to $\frac{1}{\alpha_p}$, and the number S^* of families is of the same order as the size of the market M.

Proof. In a symmetric equilibrium, $p_j^* = p^*$ and $N_j^* = N^*$. Thus, $s_j^* = 1/S^*$, and we can rewrite the first-order conditions as

$$p^* \alpha_p \left(1 - \frac{1}{S^*}\right) = 1, \tag{3}$$

$$p^* M \alpha_N \left(1 - \frac{1}{S^*}\right) = FN^*S^*, \tag{4}$$

and the free-entry condition as

$$\pi(N^*, p^*) = \frac{p^* M}{S^*} - K - FN^* = 0. \tag{5}$$

We can substitute equations (3) and (4) into the zero-profit condition (5), rearrange, and obtain

$$M = \frac{\alpha_p K S^*(S^* - 1)}{\alpha_N + S^*(1 - \alpha_N)}. \tag{6}$$

(i) Consider a large market—that is, $M \to +\infty$—and suppose $S^* \to +\infty$. Then, the left-hand side of equation (6) diverges to positive infinity, whereas the right-hand side of equation (6)

diverges to negative infinity because its sign is equal to the sign of $\frac{\alpha_p K}{1-\alpha_N} < 0$. However, this is impossible. Thus, the only possibility is $\alpha_N + S^*(1-\alpha_N) = 0$, which we can solve for S^* as

$$S^* = \frac{\alpha_N}{\alpha_N - 1}.$$

Substituting $S^* = \frac{\alpha_N}{\alpha_N - 1}$ into equations (3) and (4), we obtain that the price p^* is equal to $\frac{\alpha_N}{\alpha_p}$ and the ratio $\frac{N^*}{M}$ is equal to $\frac{(\alpha_N - 1)}{\alpha_p F}$.

(ii) In a large market—that is, $M \to +\infty$—the left-hand side of equation (6) diverges to positive infinity and the right-hand side has a positive sign. Thus, S^* has to diverge to positive infinity as well. Moreover, letting $M \to +\infty$ and $S^* \to +\infty$ in equations (3) and (4), we obtain that the price p^* is equal to $\frac{1}{\alpha_p}$ and the number N^* of funds is equal to $\frac{\alpha_N K}{(1-\alpha_N)F}$. Substituting $p^* = \frac{1}{\alpha_p}$ and $N^* = \frac{\alpha_N K}{(1-\alpha_N)F}$ into equation (4), and noting that $\frac{1}{S^*} = 0$, we obtain

$$\frac{S^*}{M} = \frac{(1-\alpha_N)}{\alpha_p K}.$$

Proposition 1 highlights the stark effects of demand-side spillovers on firms' equilibrium strategies. In particular, two distinct types of equilibria arise, depending on whether the parameter α_N is larger or smaller than one—that is, the rate at which fixed costs increase as families add new funds. If demand spillovers are strong ($\alpha_N > 1$), competition induces families to offer a large number of funds rather than low fees. As a result, the industry is concentrated even in large markets. However, if demand spillovers are weak ($\alpha_N \leq 1$), the number of funds that each family offers is limited, fees are lower, and the industry is more fragmented in larger markets. In particular, the number S^* of families converges to a finite constant when spillovers are strong; instead, S^* is of the same order as the market size M when spillovers are weak, implying that the number of families is larger if spillovers are weak. Figure 2 further displays equilibrium outcomes in markets of moderate sizes—obtained solving the system of equations (3)–(5)—confirming the comparisons of Proposition 1.

The intuition for the results of Proposition 1 closely follows the arguments of Shaked and Sutton (1987). When decision makers place a high value on the number of funds—that is, the dimension of vertical differentiation between families—a family that offers more funds than its competitors can undercut their fees and gain a proportionally larger market share because offering more funds affects fixed rather than marginal costs. Thus, only a few large families can survive in equilibrium, and the industry stays concentrated even in a large market. When, instead, decision makers place a relatively low value on the number of funds ($\alpha_N \leq 1$), increasing its funds offerings and undercutting rivals' fees are not profitable, in particular when firms are already choosing low fees that almost equal marginal cost—that is, zero. Moreover, in a large market, even very low fees generate enough variable profits to cover fixed costs. Thus, many families survive in equilibrium, and the industry is fragmented.

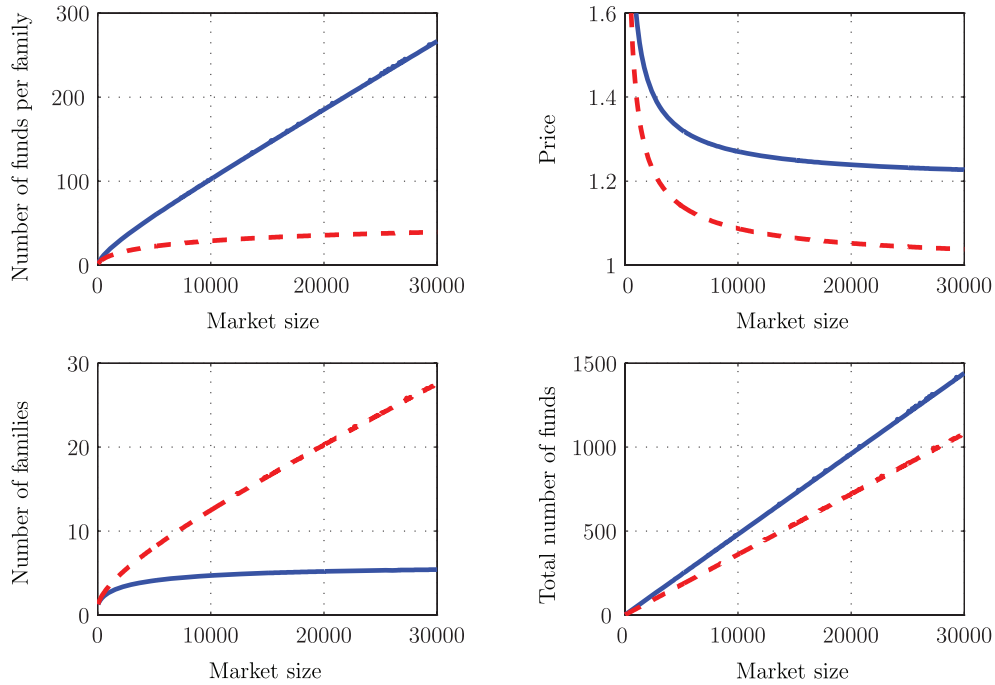
Proposition 1 suggests that strong spillovers provide a potential explanation for the limited price competition observed in the mutual fund industry (Sirri and Tufano, 1993; Khorana and Servaes, 2007; Gil-Bazo and Ruiz-Verdú, 2009). Moreover, the strength of spillovers also affects the aggregate number of funds offered to investors, providing a theoretical explanation for the existence of a large number of funds, which has been viewed as a puzzle (Massa, 1997).

Corollary 2. The total number of funds N^*S^* is higher and the market share $\frac{M}{N^*S^*}$ of each fund is lower if spillovers are higher (higher α_N).

Proof. Using the expressions for N^* and S^* derived in the proof of Proposition 1, we obtain that $N^*S^* = \frac{\alpha_N M}{\alpha_p F}$. Thus, N^*S^* is increasing and $\frac{M}{N^*S^*}$ is decreasing in α_N, the extent of demand-side spillovers.

FIGURE 2

COMPARATIVE STATISTICS



Note: This figure compares market outcomes when spillovers are weak (dashed lines) and when spillovers are strong (solid lines). The parameters are $K = 150$, $F = 25$, $\alpha_p = 1$, and $\alpha_N = 1.2$ (solid lines) or $\alpha_N = 0.9$ (dashed lines).

In summary, although the life cycle of the mutual fund industry, like that of any other industry, is clearly more complex than our simple theoretical framework, the model can explain salient features of competition among fund families. Furthermore, the analysis of Shaked and Sutton (1987) and Sutton (1991) indicates that the substance of the results does not depend on some of the simplifying assumptions imposed. In particular, the results are robust to several important extensions: (i) heterogeneous preference α_N across consumers; (ii) additional horizontal or vertical characteristics that affect investors' utility function; (iii) heterogeneous fixed costs F across firms; and (iv) sequential entry of firms.

5. Empirical analysis

■ This section first describes the data employed in this study. It then illustrates some striking differences between the institutional and retail segments of the mutual fund industry. In particular, in the retail segment, families offer more funds, funds have higher fees, the market is more concentrated, and the total number of funds is larger than in the institutional segment. These patterns are exactly what Proposition 1 and Corollary 2 imply if spillovers are stronger in the retail segment than in the institutional segment. Thus, the following subsections measure and compare demand spillovers in the two segments.

□ **Data.** The empirical analysis is based on the mutual fund database compiled by the CRSP. This data set provides mutual fund data for all funds, including defunct funds. We include in the sample all active funds for which the main fund characteristics—total net assets, returns, turnover ratio, expense ratio, load, and so forth—and their family identifier are available.

CRSP also reports the zip code of the headquarters of each fund family. We use this information to complement the CRSP Mutual Fund Database with data from the County Business Patterns (CBP).[8] Specifically, we match the zip code of the headquarters of each family to the corresponding county in CBP. The CBP data set is unique in its coverage of all private sectors of the economy and its link to location. In particular, for each county and four-digit industry pair, the CBP reports total number of plants, total employment, and total payroll.

Because CRSP reports family headquarters' zip codes starting from 1999, we restrict our regression analysis to the period from 1999 to 2007, although we present some general industry trends for the period 1992–2007.

□ **Variables and summary statistics.** We follow the CRSP Mutual Fund Database and previous literature for the construction of most of the variables. CRSP treats multiple share classes offered by a fund as separate entities, and we define each single share class as a separate fund.[9] (Because this definition may be arbitrary, we also perform our analysis using an alternative classification that does not depend on this definition; see below.) CRSP classifies each fund as retail or institutional, and we follow this definition. We then treat retail and institutional funds as separate segments. In particular, we include only funds within each segment when we calculate aggregate variables.

From CRSP, we obtain Fund Tna (fund's total net assets), and we construct Fund Market Share$_{jt}$, the market share of fund j in year t, as $\frac{100*\text{Fund Tna}_{jt}}{\sum_j \text{Fund Tna}_{jt}}$, where $\sum_j$Fund Tna$_{jt}$ is computed summing over retail or institutional funds only. CRSP reports each fund's monthly return r_{jikmt}, and we construct a Fund Return as $R_{jikt} = \prod_{m=1}^{12}(1 + r_{jikmt})$, where j denotes a fund, i denotes a family, k denotes a category/investment objective (i.e., money market, U.S. equities, etc.), m denotes a month, and t denotes a year. Fund Fees are equal to regular expenses, plus one seventh of front- and rear-end fees charged by the family. Hence, we are assuming a seven year investment horizon, as in previous studies (Sirri and Tufano, 1998; Khorana and Servaes, 2007). Fund Turnover follows the CRSP definition.

We construct most family-level variables by averaging the fund-level variables of all funds within a family and within each segment (retail or institutional), weighting each fund by its assets. Moreover, in order to partially take into account the heterogeneity in family strategies across different categories of funds, we construct some family-level variables by taking the weighted average of fund characteristics calculated as deviations from the average characteristics of all funds in the same category. For example, the aggregate performance of family i in year t is measured as the Family Excess Return:

$$\text{Family Excess Return}_{it} = \sum_j \omega_{jikt}\left(R_{jikt} - \sum_l \omega_{l\cdot kt} R_{l\cdot kt}\right),$$

where, as before, j denotes a fund, i denotes a family, k denotes a category/investment objective, and t denotes a year. ω_{jikt} is fund j's share of family i's total assets in year t; R_{jikt} is the return of fund j of family i belonging to category k in year t; and $\omega_{l\cdot kt}$ is fund l's share of category k's assets in year t. Thus, $\sum_l \omega_{l\cdot kt} R_{l\cdot kt}$ is the weighted average return of all funds belonging to category k in year t, and $R_{jikt} - \sum_l \omega_{l\cdot kt} R_{l\cdot kt}$ is the year t -abnormal return of fund j of family i belonging to category k. Thus, Family Excess Return$_{it}$ adjusts for all factors that may affect all funds in the same investment category (Khorana and Servaes, 2007). Moreover, we also construct

[8] The CBP is available at www.census.gov/econ/cbp/index.html.

[9] Fund sponsors frequently offer different share classes of a single portfolio to investors, primarily load or no-load classes. The three share classes commonly offered by multiple-class funds are denoted A, B, and C. The A class is the traditional class in which investors pay a front-end load and an annual 12b-1 fee of 25–35 basis points to compensate brokers. In comparison, the B and C classes have no front-end loads but may charge a contingent deferred sales load upon exit and usually charge higher annual 12b-1 fees of about 1%. See Nanda, Wang, and Zheng (2009) for an analysis of funds' decision to issue multiple share classes.

the variable Family Variance Excess Return as $\sum_j \omega_{jikt}(R_{jikt} - \sum_l \omega_{l\cdot kt} R_{l\cdot kt})^2$ to control for the heterogeneity of fund returns within a family. Similarly, we construct the variable Family Fees and Family Turnover as $\sum_j \omega_{jikt} x_{jikt}$, where x_{jikt} is the fee or the turnover, respectively, of fund j of family i belonging to category k in year t.

The main explanatory variables of interest measure the number of products offered by a family in the corresponding segment (retail or institutional), and we measure it in two complementary ways. The first one—Family Funds—is the total number of funds offered by a family. Because we are treating each share class offered by a fund as a separate entity, Family Funds may, perhaps, overestimate the number of funds available to investors. Thus, we also construct a second measure that is not affected by this potential mismeasurement: Family Categories, the total number of categories/investment objectives in which a family offers at least one fund. Hence, Family Categories performs an important robustness check (which we could have performed by defining a fund as the aggregation of all share classes of a unique portfolio), and has the additional advantage of providing a measure of families' offerings that emphasizes product-line breadth across different investment objectives. CRSP offers several distinct classifications of funds' categories/investment objectives. We employ the most detailed classification available, the Lipper objective codes classification. Thus, the variable Family Categories is a count of the number of distinct Lipper objective codes offered by a fund family.

Finally, we obtain from the CBP some variables that we use as instruments in our empirical analysis: Total Establishments and Employment Financial Sector are defined exactly as in the CBP, where the financial sector refers to either SIC 60 or NAICS 52; Wage is calculated as total payroll divided by total employment, and Wage Financial Sector is total payroll in the financial sector divided by Employment Financial Sector. All these variables refer to the county in which the zip code of the family's address reported in CRSP is located. Moreover, we also calculate the distance of the family headquarters from New York City as $d_i = \sqrt{(lat_i - lat_{NY})^2 + (long_i - long_{NY})^2}$, where lat_j and $long_j$ are the latitude and longitude of zip code j. Zip code 10012 is used for all families located in New York City.

Table 1 provides summary statistics of the main variables used in the empirical analysis. The first two columns of data refer to the retail segment of the market. There are 66,270 fund-year observations in this segment. On average, a retail fund has a market share of approximately 1%, and there is considerable heterogeneity in the size of funds: the standard deviation of the market share is equal to 5.5%. The (unweighted) average fee in the retail sector equals 1.92% and the average return is 7.6%. The average retail fund belongs to a family with 120 funds spanning 27 different categories. The heterogeneity across families is substantial: the standard deviations of Family Funds and Family Categories are equal to 93 funds and 16 categories, and all other family variables exhibit significant variations across families.

The last two columns present summary statistics for the institutional segment of the market. As Section 3 argues, demand spillovers should not play the same role in the two segments of the market. Indeed, Table 1 presents some suggestive patterns. The average institutional fund has a larger market share, is cheaper and younger, yields a higher return, and has a lower turnover than a retail fund. Interestingly, the average institutional fund belongs to a smaller family than a retail fund does—a family with 60 funds spanning 18 different categories.

□ **A graphical comparison.** Figure 3 displays striking differences between the institutional and retail segments of the market over the period 1992–2007. First, the top left plot shows that families offer a larger (asset-weighted) number of funds in the retail segment than in the institutional segment. Second, the top right plot shows that (asset-weighted) fees are higher in the retail segment. Third, the bottom left plot shows that the retail segment is less concentrated than the institutional segment: the market share of the largest four families of institutional funds is about only half of the market share of the largest four families of retail funds. Moreover, whereas the retail segment has become *more* concentrated over time, the institutional segment has become

TABLE 1 Summary Statistics

	Retail Funds		Institutional Funds	
	Mean	Standard Deviation	Mean	Standard Deviation
Fund Market Share	.0106	.0558	.0207	.0993
Fund Fees	.0192	.0080	.0099	.0046
Fund Age	9.043	8.695	6.868	4.936
Fund Return	1.076	.1782	1.0869	.1674
Fund Turnover	1.044	2.022	.9587	1.129
Family Funds	120.0	93.71	60.57	58.84
Family Categories	26.8	15.90	18.54	9.90
Family Excess Return	−.0050	.0670	−.0003	.0437
Family Variance Excess Return	.0117	.0706	.0059	.0233
Family Turnover	.8463	.8601	.8540	.6097
Employment Financial Sector	122,397	117, 365	121,274	113,971
Total Establishments	53,661	48,473	59,889	59,373
Wage ($1000)	54.8	18.0	53.8	18.6
Wage Financial Sector ($1000)	109.4	60.9	108.5	63.4
Distance from New York	11.4	15.1	13.3	16.8
No. Observations	66,271		22,524	

Notes: This table provides summary statistics of the variables used in the empirical analysis. Most fund variables follow the exact definition of the CRSP Mutual Fund Database. CRSP classifies each fund as a retail or institutional fund, and we follow this definition. Fund Market Share is the market share of a fund in a given year, equal to $\frac{100*FUND\,TNA_{jt}}{\sum_j FUND\,TNA_{jt}}$, where Fund Tna is the fund's total net assets. Fund Fees are equal to regular expenses, plus one seventh of front- and rear-end fees charged by the family. Fund Age is the number of years since the fund was established. Fund Return is the annual return of the fund. Fund Turnover is the fund turnover ratio, expressed as a percentage of the fund. Family Funds is the number of funds offered by the fund family. Family Categories is the number of Lipper categories in which the fund family offers at least one fund. Family Excess Return is the asset-weighted, objective-adjusted average family abnormal return. Family Variance of Excess Return is the asset-weighted, objective-adjusted variance of family abnormal returns. Family Turnover is the asset-weighted average turnover across all the family's funds. Total Establishments is the number of establishments in the county corresponding to the zip code of the family address reported in CRSP. Employment Financial Sector is the total employment in the financial sector in the county corresponding to the zip code of the family address reported in CRSP. Wage is total payroll divided by total employment in the county corresponding to the zip code of the family address reported in CRSP. Wage Financial Sector is total payroll in the financial sector divided by Employment Financial Sector in the county corresponding to the zip code of the family address reported in CRSP. These last four variables are obtained from County Business Patterns. Distance from New York is the distance of the family address from New York City. All data refer to the years from 1999 to 2007.

less concentrated. Fourth, the bottom right panel documents that the total number of funds is larger in the retail than in the institutional segment.

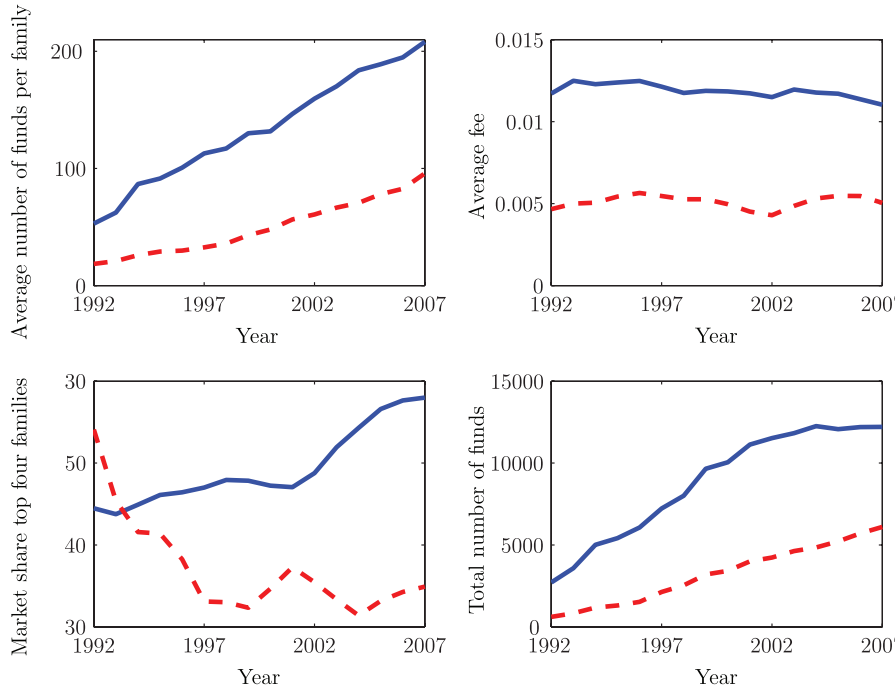
These patterns match exactly the predictions of the simple model presented in Section 4 if spillovers are stronger in the retail segment than in the institutional segment. Thus, the goal of the next subsections is to measure these spillovers by estimating a demand system in each segment, enriching the demand system given by equation (1) of the theoretical model.[10]

□ **Demand spillovers in the retail segment.** We proceed in three complementary ways to measure demand spillovers in the retail segment of the industry. First, we investigate whether a family that offers more funds or categories than its rivals has a proportionally larger market share (and assets under management). Second, we investigate whether a fund whose family offers more funds or categories than its rivals has a larger market share. Third, we investigate whether a family that offers more funds or categories than its rivals has a larger market share in the categories that it offers.

[10] An alternative empirical design may be to perform one family-level regression and one fund-level regression, using dummy variables to distinguish the key coefficients of retail funds from institutional funds. Although this is certainly feasible, the main advantage of the empirical design followed below is that it does not constrain the coefficients of the additional control variables to be the same in the retail and institutional segments.

FIGURE 3

A GRAPHICAL COMPARISON: INSTITUTIONAL FUNDS VERSUS RETAIL FUNDS



Notes: This figure presents several comparisons between Institutional Funds (dashed lines) and Retail Funds (solid lines). The top left plot displays the asset-weighted average number of funds per family. The top right plot displays the asset-weighted average fee. The bottom left plot displays the market share of the largest four families. The bottom right plot displays the total number of funds. All plots correspond to the period 1992–2007.

Family-level evidence. To investigate the importance of demand spillovers at the family level, we estimate a demand system rearranging the market share equation (1) of the theoretical model (see Berry, 1994), also including additional variables that control for factors that the model does not consider. Specifically, we estimate the following regression equation[11]:

$$\text{Log(Family Market Share}_{it}) = \alpha \text{Log(Variety}_{it}) + \gamma Z_{it} + \eta_t + \epsilon_{it}. \tag{7}$$

The dependent variable Log(Family Market Share$_{it}$) is the log of a family's aggregate market share—that is, the total net assets of family i in year t divided by the total net assets of all families in year t. The key variable of interest is Variety$_{it}$, defined as either Family Funds—that is, the total number of funds offered by a family—or Family Categories—that is, the total number of categories/investment objectives in which a family offers at least one fund—offered by family i in year t. Z_{it} is a vector of variables specific to each individual family, such as the average return of all funds, the variance of returns of all funds, and the (log of the) age of the family, measured as the age of the oldest active fund. η_t is a year fixed effect that captures aggregate trends in the industry (e.g., the growth of ETFs or the inflows by new, and perhaps less sophisticated, households). ϵ_{it} is an idiosyncratic unobserved component. Because the regression equation (7) includes year dummies, the estimated coefficients are identical if Log(Family Tna$_{it}$) is the dependent variable.

The use of the variables Family Funds and Family Categories creates a potential challenge to identifying the effects of the demand spillovers in equation (7). The reason is that both measures of varieties/product-line breadth are chosen by each profit-maximizing family and, thus, are

[11] More precisely, equation (7) obtains after taking logs in equation (1), recognizing that the log of the denominator of the right-hand side is constant within a year. Thus, the year fixed effect in equation (7) captures it.

endogenous. In particular, any observed correlation between our measures of a family's varieties and a family's aggregate market share could be due to either demand or supply reasons, or both. The literature argues that two contrasting effects may apply. First, if supply-side economies of scale—that is, lower research, product development, marketing, advertising, and various administrative costs—and economies of scope—that is, cost complementarities—are significant, larger fund families may enjoy lower costs per fund and, thus, offer more funds (Khorana and Servaes, 1999). Second, families may launch new funds if the current market shares are low for observed and unobserved reasons. Specifically, Ippolito (1992) and Khorana and Servaes (1999) argue that a family may open new funds to disguise the poor performance of existing funds, and Nanda, Wang, and Zheng (2004) show that families, especially lower-performance families, introduce multiple funds to maximize the chances of having a fund with superior returns—that is, a "star" fund. The first effect should bias the OLS coefficients of product offerings upward, whereas the second effect should bias the OLS coefficients of product offerings downward. As a result of these two contrasting effects, the sign of OLS estimates' bias is ambiguous.

Thus, the previous *supply-side*arguments could invalidate the tests of our predictions. As in a standard demand and supply estimation, in order to identify *demand-side* spillovers, we employ variables that shift Family Funds and Family Categories for *supply-side* reasons. Specifically, an influential economic geography literature has empirically documented the spatial agglomeration of firms belonging to the same industry (for a summary, see Fujita, Krugman, and Venables, 1999), and a few articles have shown that firms located in these areas where an industry concentrates are, on average, larger than firms in the same industry outside such areas (for evidence across industries, see Holmes and Stevens, 2002; for evidence on the mutual fund industry, see Christoffersen and Sarkissian, 2009). The general reason is that firms located where an industry concentrates enjoy larger labor pools and larger flows of ideas (Marshall, 1920), thereby growing in size. Indeed, applying these insights to the mutual fund industry, Christoffersen and Sarkissian (2009) emphasize the difficulty of hiring fund managers away from financial centers and the positive information spillovers of these financial centers. Hence, everything else equal, the *supply* of funds/categories should be higher if a family is located where employment in the financial sector is higher. Thus, we retrieve from CRSP the zip code of the headquarters of each family, and we match it to Employment Financial Sector (and to Total Establishments) of the corresponding county to obtain a measure of the density of the financial sector, as well as relative to other industries. Moreover, we further employ as an instrument Log(Distance from NYC), the distance of each family's headquarters from New York City. Table A1 reports the "first-stage" regressions of the endogenous variables Family Funds and Family Categories. The key findings are that, on average, families located in counties with a larger employment in the financial sector (relative to the total number of establishments) offer a larger number of funds or funds in a larger number of categories.

The validity of these instruments relies on two key assumptions: (i) each fund family represents a small fraction of Employment Financial Sector in the cross-section of U.S. counties; and (ii) the exclusion restriction that, for example, Employment Financial Sector in Los Angeles or Boston does not *directly* affect the demand and the market share of American Funds or Fidelity, whose headquarters are in Los Angeles and Boston, respectively.[12] Because fund families tend to locate in large financial centers such as New York, San Francisco, and Chicago (Christoffersen and Sarkissian, 2009), while competing to attract investors nationwide, these assumptions seem reasonable features of the mutual fund industry. Nonetheless, here we check the robustness of our results against potential concerns about the validity of these assumptions and, thus, of the instruments. In addition, it is important to note that the thrust of our empirical analysis relies on the comparison between spillovers in the retail and institutional segments of the market. Our

[12] The instruments may also affect the market share of the family if, for example, higher wages attract more skilled managers who, in turn, generate higher returns. All our regressions include returns, thus controlling directly for this effect.

TABLE 2 Demand Spillovers: Family Regressions, Retail Segment

	(1)	(2)	(3)	(4)
Log(Family Funds)	1.3773		1.4816	
	(.0335)		(.2003)	
Log(Family Categories)		1.8233		2.0126
		(.0538)		(.2962)
Log(Family Fees)	−1.5756	−1.2597	−2.9583	−2.6735
	(.1333)	(.1540)	(.5871)	(.6467)
Family Age	.0310	.0228	.0330	.0248
	(.0082)	(.0082)	(.0092)	(.0089)
Family Excess Return	1.5606	1.6315	.9222	.9911
	(.1534)	(.1643)	(.3150)	(.3397)
Family Variance Excess Return	−.3701	−.3868	.2458	.2581
	(.1972)	(.1985)	(.3139)	(.3244)
Family Turnover	−.0581	−.0583	.0143	.0128
	(.0333)	(.0355)	(.0394)	(.0412)
Year Fixed Effects	Yes	Yes	Yes	Yes
No. Observations	3,753	3,753	3,753	3,753
No. Families	528	528	528	528

Notes: This table reports the results of OLS and IV regressions that investigate whether a family that offers more funds or categories than its rivals in the retail segment has a proportionally larger market share in the retail segment. Columns (1) and (2) refer to OLS regressions, and columns (3) and (4) refer to IV regressions. Each observation corresponds to a family-year pair. The dependent variable is Log(Family Market Share), the log of a family's aggregate market share. All independent variables are defined in Section 5. In the IV specifications (3) and (4), the endogenous variables Family Funds, Family Categories, and Family Fees are instrumented using the instruments described in the text. All regressions further include year fixed effects. The data refer to the years from 1999 to 2007. Standard errors in parentheses are clustered at the fund family level.

results will differ across these two segments, and reasonable arguments against the validity of the instruments should account for this difference.

Furthermore, we also instrument Family Fees—the other endogenous variable in the model—using instruments introduced by Berry, Levinsohn, and Pakes (1995) to estimate utility functions with unobservable characteristics: summary measures of other families' characteristics in the current year. In particular, we use average returns and variance of returns. These instruments capture the effect of a family's relative position in characteristics space, assumed to be exogenous, on their endogenous decisions, independent of the family's unobservable characteristics.

Columns (1) and (2) of Table 2 present OLS estimates of the coefficients of equation (7), and columns (3) and (4) present IV estimates. The coefficients of Log(Family Funds) and Log(Family Categories) are not statistically different between the OLS and IV specifications, indicating that the endogeneity of offerings may not affect the estimates of these coefficients. Instead, the IV estimates of the coefficients of Log(Family Fees) are lower and statistically different from the OLS estimates, confirming the usual upward bias of OLS estimates due to the positive correlation between unobserved characteristics and prices.

Consistent with our demand spillover arguments, the positive coefficients of Log(Family Funds) and Log(Family Categories) indicate that investors are more likely to allocate their assets in families that offer more funds. However, strictly speaking, the positive coefficients of Log(Family Funds) and Log(Family Categories) do not necessarily imply demand spillovers. For example, if, for whatever reason, investors allocate their money randomly across funds, then, mechanically, families that offer more funds would have a higher market share. Nevertheless, the coefficients are estimated to be statistically larger than one, indicating that families that offer more funds (categories) have a proportionally larger market share. The magnitude of the effects is also nontrivial: according to the IV estimates of columns (3) and (4), a family that offers 10% more funds has a 14.8% higher market share, and a family that offers 10% more categories has a 20.1% higher market share.

Fund-level evidence. To investigate the importance of demand-side spillovers at the fund level, we enrich equation (7) as follows:

$$\text{Log(Fund Market Share}_{jkit}) = \alpha \text{Log(Variety}_{it}) + \beta X_{jkit} + \gamma Z_{it} + \zeta_{kt} + \epsilon_{jkit}. \quad (8)$$

The dependent variable Log(Fund Market Share$_{jkit}$) is the log of the market share of fund j in category k belonging to family i in year t. Variety$_{it}$ is either Family Funds or Family Categories offered by family i in year t. X_{jkit} is a vector of variables specific to each individual fund, such as the annual return of the fund, the (log of the) age of the fund, and the turnover. Z_{it} is a vector of variables specific to each individual family, such as the average return of all funds, the variance of returns of all funds, and the (log of the) age of the family, measured as the age of the oldest active fund. ζ_{kt} is a category-year fixed effect and ϵ_{jkit} is an idiosyncratic unobserved component. Hence, equation (8) is similar to equation (7), but it uses a richer set of characteristics at the individual fund level, thus more carefully controlling funds' observable heterogeneity across families.

Specifications (1) and (2) in Table 3 report OLS estimates of the coefficients and specifications (3)–(10) report IV estimates, as the endogenous variables Family Funds or Family Categories are instrumented with the same *supply-side* shifters described in the previous section. Similarly, in specifications (3)–(10), we instrument Fund Fees using the average excess return and average variance of excess returns of rival families, as well as average returns of other funds operated by rival families in the same category. Specifications (1)–(4) do not include category fixed effects; specifications (5) and (6) include them; and specifications (7) and (8) include category-year fixed effects. The coefficients of the two variables capturing the family variety of funds—Log(Family Funds) and Log(Family Categories)—are significantly higher than zero in all specifications. (As in the previous analysis at the family level reported in Table 2, the coefficients of Log(Family Funds) and Log(Family Categories) are not statistically different between the OLS and IV specifications.) The signs of these coefficients indicate that, on average, the market share of a retail fund is larger if the number of funds (categories) offered by its family is larger. This is a strong confirmation of the results of Table 2 and reinforces the idea that demand spillovers play an important role in the retail segment of the market. For example, according to the coefficients of specifications (7) and (8), a family that offers 10% more funds than its rivals has a 3.7% higher market share for each fund of the same family, and a family that offers 10% more categories than its rivals has a 5.3% higher market share for each fund of the same family. Furthermore, the coefficients of Family Excess Return are positive in all specifications, although sometimes not significant, confirming the importance of within-family performance spillovers as in Sirri and Tufano (1998), Nanda, Wang, and Zheng (2004), and Ivković (2004).

Columns (9) and (10) of Table 3 report the estimates of the coefficients of equation (8) obtained on a sample of S&P 500 Index funds only. These regressions are particularly interesting because the sample is restricted to (almost) homogeneous products that passively follow an index. Thus, we should expect the effect of demand spillovers on this sample of S&P 500 Index funds to be greater than on the entire sample, for two main reasons. First, fund heterogeneity (observed and unobserved) should play a limited role in index funds. Second, an S&P 500 Index fund may be a "focal" fund in which investors "park" their assets when they unload them from other funds within the family. Indeed, the coefficients reported in columns (9) and (10) exactly indicate that the effect of spillovers is greater on this sample of S&P 500 Index funds: the coefficients of the two variables capturing the family variety of funds—Log(Family Funds) and Log(Family Categories)—are significantly larger in columns (9) and (10) than in columns (7) and (8), confirming the importance of demand spillovers. More precisely, the market share of the S&P 500 Index fund offered by a family that offers 10% more funds than its rivals is 7.6% higher, and the market share of the S&P 500 Index fund offered by a family that offers 10% more categories than its rivals is 14.6% higher.

Category-level evidence. We now investigate the importance of demand-side spillovers by examining how the number of family products affects the family's market share in each Lipper

TABLE 3 Demand Spillovers: Fund Regressions, Retail Segment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log(Family Funds)	.2301		.2479		.4087		.3757		.7692	
	(.0465)		(.0936)		(.1060)		(.1077)		(.4042)	
Log(Family Categories)		.3316		.3170		.6008		.5345		1.4686
		(.0685)		(.1365)		(.1608)		(.1622)		(.7969)
Log(Fund Fees)	−1.2625	−1.2134	−1.8956	−1.8237	−4.0068	−3.8006	−2.2786	−2.0771	−1.2280	−1.2394
	(.1185)	(.1164)	(.2639)	(.2628)	(.6715)	(.6625)	(.8219)	(.7934)	(.3137)	(.3103)
Fund Return	1.2449	1.2363	1.3359	1.3239	.6554	.6524	1.0942	1.1086		
	(.0981)	(.0982)	(.1043)	(.1040)	(.0636)	(.0635)	(.1441)	(.1413)		
Fund Turnover	−.0551	−.0512	−.0480	−.0428	−.0518	−.0459	−.0630	−.0555		
	(.0130)	(.0123)	(.0118)	(.0105)	(.0118)	(.0121)	(.0112)	(.0117)		
Fund Age	.0915	.0909	.0830	.0823	.0724	.0711	.0866	.0861	.1753	.1504
	(.0051)	(.0052)	(.0050)	(.0051)	(.0065)	(.0067)	(.0077)	(.0078)	(.0325)	(.0364)
Family Excess Return	1.7197	1.7230	1.2196	1.2279	.3956	.4354	1.1267	1.1651	2.2998	2.5717
	(.3434)	(.3413)	(.3058)	(.3053)	(.4213)	(.4203)	(.4698)	(.4637)	(1.9101)	(2.1219)
Family Variance Excess Return	−.4164	−.3781	−.2237	−.2350	.2569	.2531	−.1697	−.1909	−3.0000	−1.3135
	(.2803)	(.2753)	(.2664)	(.2448)	(.2433)	(.2403)	(.2717)	(.2705)	(3.9430)	(4.9756)
Family Turnover	−.0061	.0015	.0453	.0475	.0428	.0587	.0096	.0218	−.9059	−.8483
	(.0536)	(.0543)	(.0536)	(.0550)	(.0610)	(.0632)	(.0605)	(.0630)	(.4624)	(.4671)
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	N/A	N/A	Yes	Yes
Category Fixed Effects	No	No	No	No	Yes	Yes	N/A	N/A	N/A	N/A
Category-Year Fixed Effects	No	No	No	No	No	No	Yes	Yes	N/A	N/A
No. Observations	66,271	66,271	66,271	66,271	66,271	66,271	66,271	66,271	497	497
No. Funds	10,768	10,768	10,768	10,768	10,768	10,768	10,768	10,768	78	78

Notes: This table reports the results of OLS and IV regressions that investigate whether a retail fund whose family offers more funds or categories than its rivals in the retail segment has a larger market share. Specifications (1) and (2) refer to OLS regressions, and specifications (3)–(10) refer to IV regressions. Specifications (1)–(8) employ the sample of all retail mutual funds. Specifications (9) and (10) employ the sample of all retail S&P 500 index funds. Each observation corresponds to a fund-year pair. The dependent variable is Log(Fund Market Share), the log of a fund's market share, in all specifications. The independent variables are defined in Section 5. The endogenous variables Family Funds, Family Categories, and Fund Fees are instrumented using the instruments described in the text in IV regressions (3)–(10). All specifications also include year fixed effects; specifications (5) and (6) further include Lipper categories fixed effects; specifications (7) and (8) include fixed effects for each Lipper category-year pair. The data refer to the years from 1999 to 2007. Standard errors in parentheses are clustered at the fund family level. N/A, not applicable.

objective code. These regressions serve two main purposes. (i) Perform a robustness check. Because the fund-level regressions treat each share class offered by a fund as a separate entity, the results reported in Table 3 may overestimate the number of independent observations, potentially biasing some coefficients. (ii) Investigate the role of spillovers across different investment objectives, rather than across different funds.

With these goals in mind, we estimate the following regression:

$$\text{Log(Cat-Fam Market Share}_{kit}) = \alpha \text{Log(Variety}_{it}) + \beta X_{kit} + \gamma Z_{it} + \zeta_{kt} + \epsilon_{kit}. \quad (9)$$

The dependent variable is the log of the market share of all funds in category k belonging to family i in year t. Variety_{it} is either Family Funds or Family Categories offered by family i in year t. X_{kit} is a vector of variables specific to each individual family-category pair, such as the funds' average return, the variance of funds' returns, and the (log of the) age of the category, measured as the age of the oldest active fund. Z_{it} is a vector of variables specific to each individual family, such as the average return of all funds, the variance of returns of all funds, and the (log of the) age of the family, measured as the age of the oldest active fund. ζ_{kt} is a category-year fixed effect and ϵ_{kit} is an idiosyncratic unobserved component.

Columns (1) and (2) in Table 4 report OLS estimates of the coefficients, and columns (3)–(10) report IV estimates, as the endogenous variables—that is, Family Funds or Family Categories, and Family-Category Fees—are instrumented with the *supply-side* shifters used in the fund-level regressions reported in previous sections. Columns (1)–(8) report the results for all retail category-family pairs. Specifications (1)–(4) do not include category fixed effects; specifications (5) and (6) include them; and specifications (7) and (8) include fixed effects for each category-year pair. The coefficients of the two variables capturing the family variety of funds—Log(Family Funds) and Log(Family Categories)—are significantly higher than zero in all specifications. The signs of these coefficients indicate that, on average, the market share of a category of funds is larger if the number of funds (categories) offered by its family is larger. This confirms the results of Tables 2 and 3 that demand spillovers play an important role in the retail segment of the market, and indicates that those results were not an artifact of measurement error. In particular, the coefficients of specifications (7) and (8) imply that a family that offers 10% more funds than its rivals has a 6.3% higher market share for each category of funds of the same family, and a family that offers 10% more categories than its rivals has a 9% higher market share for each category of funds of the same family. As in the previous family- and fund-level analyses, the coefficients of Log(Family Funds) and Log(Family Categories) are not statistically different between OLS and IV specifications.

Columns (9) and (10) of Table 4 report the estimates of the coefficients of equation (9) obtained on the category of S&P 500 Index funds only. In these specifications, we use our supply shifters only to instrument the endogenous variables Family Funds or Family Categories, because instrumenting Family Fees generates unstable results. With this caveat in mind, the coefficients reported in columns (9) and (10) confirm that the effect of spillovers is greater on this sample of S&P 500 Index funds: the coefficients of the two variables capturing the family variety of funds—Log(Family Funds) and Log(Family Categories)—are larger in columns (9) and (10) than in columns (7) and (8), confirming the importance of demand spillovers. More precisely, the market share of the S&P 500 Index category of funds offered by a family that offers 10% more funds than its rivals is 11% higher, and the market share of the S&P 500 Index category of funds offered by a family that offers 10% more categories than its rivals is 22% higher.

□ **Demand spillovers in the institutional segment.** We now measure the magnitude of demand spillovers in the institutional segment of the market and compare it with the magnitudes estimated previously in the retail segment. This comparison is useful to understanding the differential patterns between the institutional and retail segments highlighted in Figure 3. In particular, Proposition 1 and Corollary 2 of our simple model suggest that these differential

TABLE 4 Demand Spillovers: Category Regressions, Retail Segment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log(Family Funds)	.4508		.4340		.6418		.6328		1.1477	
	(.0335)		(.0862)		(.1437)		(.1381)		(.5670)	
Log(Family Categories)		.6080		.6353		.9491		.9063		2.2886
		(.0502)		(.1338)		(.2180)		(.2117)		(1.0852)
Log(Category-Family Fees)	−.7774	−.6730	−1.0530	−.9280	−1.5045	−1.3351	−1.6820	−1.3202	−.8889	−.9100
	(.1059)	(.1112)	(.3288)	(.3331)	(.6750)	(.6975)	(.6399)	(.6615)	(.2816)	(.2485)
Category-Family Return	1.5237	1.5107	1.5624	1.5415	.9452	.9629	1.4146	1.4099		
	(.0766)	(.0782)	(.1071)	(.1086)	(.0759)	(.0770)	(.0759)	(.1615)		
Category-Family Turnover	−.0310	−.0232	−.0324	−.0261	−.0421	−.0331	−.0404	−.0306		
	(.0109)	(.0110)	(.0105)	(.0106)	(.0114)	(.0116)	(.0113)	(.0115)		
Category-Family Age	.0625	.0642	.0604	.0601	.0538	.0531	.0555	.0553	.1189	.0608
	(.0043)	(.0045)	(.0048)	(.0048)	(.0068)	(.0067)	(.0064)	(.0064)	(.0595)	(.0741)
Family Excess Return	.9788	1.0350	.6460	.6920	1.0464	1.0603	.4889	.6087	6.0218	6.6240
	(.2223)	(.2317)	(.2944)	(.2985)	(.3044)	(.3089)	(.3462)	(.3489)	(2.3719)	(2.6481)
Family Variance Excess Return	−.6648	−.6722	−.5217	−.4854	−.3207	−.2441	−.3704	−.3765	−12.1359	−9.9457
	(.1776)	(.1893)	(.2319)	(.2470)	(.3092)	(.3320)	(.2943)	(.3198)	(5.1826)	(5.8704)
Family Turnover	.0039	.0083	.0191	.0317	−.0067	.0126	−.0073	.0003	−1.3496	−1.4084
	(.0347)	(.0378)	(.0408)	(.0425)	(.0480)	(.0511)	(.0456)	(.0487)	(.8537)	(.9225)
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	N/A	N/A	Yes	Yes
Category Fixed Effects	No	No	No	No	Yes	Yes	N/A	N/A	N/A	N/A
Category-Year Fixed Effects	No	No	No	No	No	No	Yes	Yes	N/A	N/A
No. Observations	23,888	23,888	23,888	23,888	23,888	23,888	23,888	23,888	285	285
No. Category-Family Pairs	3,493	3,493	3,493	3,493	3,493	3,493	3,493	3,493	39	39

Note: This table reports the results of OLS and IV regressions that investigate whether a category of retail funds offered by a family that offers more funds or categories than its rivals in the retail segment has a larger market share. Specifications (1) and (2) refer to OLS regressions, and specifications (3)–(10) refer to IV regressions. Specifications (1)–(8) employ the sample of all retail category-family pairs. Specifications (9) and (10) employ the category of retail S&P 500 index funds. Each observation corresponds to a category-family-year tuple. The dependent variable is Log(Cat-Fam Market Share), the log of a family-category's market share, in all specifications. The independent variables are defined in Section 5. In specifications (3)–(10), the endogenous variables Family Funds, Family Categories, and Category-Family Fees are instrumented using the instruments described in the text. All specifications also include year fixed effects; specifications (5) and (6) further include Lipper categories fixed effects; specifications (7) and (8) include fixed effects for each Lipper category-year pair. The data refer to the years from 1999 to 2007. Standard errors in parentheses are clustered at the fund family level.

TABLE 5 Demand Spillovers: Family Regressions, Institutional Segment

	(1)	(2)	(3)	(4)
Log(Family Funds)	1.1488		1.1540	
	(.0610)		(.4407)	
Log(Family Categories)		1.3881		1.3750
		(.0797)		(.5783)
Log(Family Fees)	−1.0018	−1.1576	−4.1425	−4.4301
	(.1714)	(.1661)	(1.5450)	(1.5682)
Family Age	.0472	.0381	.0187	.0071
	(.0138)	(.0113)	(.0259)	(.0243)
Family Excess Return	2.7677	2.7261	1.5126	1.4366
	(.6683)	(.6740)	(.5478)	(.5570)
Family Variance Excess Return	−4.4988	−4.3811	1.0796	1.2517
	(2.2674)	(2.2580)	(1.4942)	(1.5843)
Family Turnover	−.1886	−.1800	−.0355	−.0148
	(.0969)	(.0998)	(.1463)	(.1483)
Year Fixed Effects	Yes	Yes	Yes	Yes
No. Observations	1,574	1,574	1,574	1,574
No. Families	231	231	231	231

Note: This table reports the results of OLS and IV regressions that investigate whether families that offer more institutional funds or categories have a proportionally larger aggregate market share in the institutional segment. Columns (1) and (2) refer to OLS regressions, and columns (3) and (4) refer to IV regressions. Each observation corresponds to a family-year pair. The dependent variable is Log(Family Market Share), the log of a family's aggregate market share. All independent variables are defined in Section 5. In specifications (3) and (4), the endogenous variables Family Funds, Family Categories, and Family Fees are instrumented using the instruments described in the text. All regressions further include year fixed effects. The data refer to the years from 1999 to 2007. Standard errors in parentheses are clustered at the fund family level.

patterns are the natural equilibrium outcomes if demand spillovers differ in the two segments of the market.

We investigate the magnitude of demand spillovers in the institutional segment by estimating the same equations (7)–(9) on the sample of institutional funds, instrumenting the endogenous variables—that is, Family Funds or Family Categories, and Family Fees—with the same *supply-side* shifters described here: total employment in the financial sector; the total number of establishments; the distance of the headquarters from New York City; and the average excess return and average variance of excess returns of rival families.

Table 5 reports OLS and IV estimates of the coefficients of equation (7)—that is, the family-level regressions—on the sample of families of institutional funds. The dependent variable in specifications (1)–(4) is (the log of) the aggregate market share of the family—that is, Log(Family Market Share$_{it}$). OLS specifications (1) and (2) show that the coefficients of the two variables capturing the variety of families' funds—Log(Family Funds) and Log(Family Categories)—are statistically lower than the corresponding coefficients obtained on the sample of retail funds reported in Table 2. Moreover, IV specifications (3) and (4) show that we cannot statistically reject the hypothesis that the coefficients of Log(Family Funds) and Log(Family Categories) are lower than one. Thus, these IV specifications provide no evidence that families that offer more funds (categories) have a proportionally larger market share. This suggests that demand spillovers do not play an important role in the institutional segment, in sharp contrast to the evidence from the retail segment reported in Table 2.

Table 6 reports OLS and IV estimates of the coefficients of equation (8)—that is, the fund-level regressions—on the sample of institutional funds. The dependent variable in the specifications (1)–(10) is the log of each fund's market share—that is, Log(Fund Market Share$_{jkit}$). The coefficients of the two variables capturing the family variety of funds—Log(Family Funds) and Log(Family Categories)—are negative and/or not significantly different from zero. Thus, these coefficients confirm the results of Table 5: on average, the market share of (and the assets managed

TABLE 6 Demand Spillovers: Fund Regressions, Institutional Segment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log(Family Funds)	−.0894		−.2580		−.4259		−.6434		.5880	
	(.1111)		(.2069)		(.2670)		(.3018)		(.4456)	
Log(Family Categories)		−.0352		−.2560		−.6014		−.6984		.6319
		(.1073)		(.3331)		(.5363)		(.5227)		(.4060)
Log(Fund Fees)	−1.0209	−1.0138	−.8573	−.7669	−4.6212	−4.5437	−4.7364	−4.0997	−1.9303	−1.7120
	(.1110)	(.1104)	(.3810)	(.4028)	(1.8641)	(2.1978)	(1.7431)	(1.9553)	(.6410)	(.3819)
Fund Return	.8429	.8450	.6165	.5201	.3964	.3883	.0354	.1711		
	(.1877)	(.1897)	(.2492)	(.2587)	(.1166)	(.1216)	(.2660)	(.2380)		
Fund Turnover	−.0985	−.1009	−.0603	−.0683	.0899	.0850	.0913	.0660		
	(.0530)	(.0537)	(.0456)	(.0445)	(.0888)	(.1064)	(.0847)	(.0976)		
Fund Age	.1585	.1611	.1576	.1593	.1441	.1561	.1310	.1530	.1943	.1800
	(.0192)	(.0206)	(.0224)	(.0220)	(.0326)	(.0320)	(.0325)	(.0310)	(.0474)	(.0450)
Family Excess Return	1.1549	1.0166	.7631	.9190	−.9832	−1.0248	−1.1178	−.9921	−1.9303	−3.0160
	(.7585)	(.8414)	(.5670)	(.5671)	(.9095)	(1.1689)	(.9458)	(1.1599)	(.6410)	(2.1930)
Family Variance Excess Return	−.5480	−.1825	−.4352	−.5835	3.2676	3.4026	2.9026	2.7966	20.3174	8.8663
	(1.8306)	(1.8167)	(1.7760)	(1.6737)	(2.1138)	(2.3287)	(2.0922)	(2.2593)	(14.4468)	(11.4196)
Family Turnover	−.0463	−.0461	.0467	.0375	.1986	.1847	.2191	.1953	−.8458	−.9633
	(.1673)	(.1651)	(.1520)	(.1541)	(.2414)	(.2609)	(.2305)	(.2475)	(.3486)	(.2445)
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	N/A	N/A	Yes	Yes
Category Fixed Effects	No	No	No	No	Yes	Yes	N/A	N/A	N/A	N/A
Category-Year Fixed Effects	No	No	No	No	No	No	Yes	Yes	N/A	N/A
No. Observations	22,524	22,524	22,524	22,524	22,524	22,524	22,524	22,524	424	424
No. Funds	4,366	4,366	4,366	4,366	4,366	4,366	4,366	4,366	67	67

Note: This table reports the results of OLS and IV regressions that investigate whether an institutional fund whose family offers more funds or categories than its rivals in the institutional segment has a larger market share. Specifications (1) and (2) refer to OLS regressions, and specifications (3)–(10) refer to IV regressions. Specifications (1)–(8) employ the sample of all institutional mutual funds. Specifications (9) and (10) employ the sample of all retail S&P 500 Index funds. Each observation corresponds to a fund-year pair. The dependent variable is Log(Fund Market Share), the log of each fund's market share, in all specifications. The independent variables are defined in Section 5. In specifications (3)–(10), the endogenous variables Family Funds, Family Categories, and Fund Fees are instrumented using the instruments described in the text. All specifications also include year fixed effects; specifications (5) and (6) further include Lipper categories fixed effects; specifications (7) and (8) include fixed effects for each Lipper category-year pair. The data refer to the years from 1999 to 2007. Standard errors in parentheses are clustered at the fund family level.

by) an institutional fund is not higher when the number of institutional funds (categories) offered by its family is higher. This corroborates that demand spillovers do not play an important role in the institutional segment, in sharp contrast to the evidence from the retail segment reported in Table 3. Moreover, it is interesting to note that the coefficients of all other family-level variables—that is, Family Excess Return, Family Variance Excess Return, and Family Turnover—are insignificant in almost all specifications, reinforcing the idea of weaker within-family spillovers in the institutional segment than in the retail segment of the market.

Table 7 reports the estimates of the coefficients of equation (9)—that is, the category-level regressions—on the sample of institutional funds. The dependent variable in specifications (1)–(10) is (the log of) each family category's market share—that is, Log (Cat-Fam Market Share$_{jkit}$). The coefficients of the two variables capturing the family variety of funds—Log(Family Funds) and Log(Family Categories)—are again negative and/or not significantly different from zero (except in OLS specification (1)). Thus, these coefficients confirm that the family market share in a category of funds is not higher when the number of institutional funds (categories) offered by the family is higher. The difference in the signs of the coefficients and/or in their statistical (in)significance across specifications indicates that demand spillovers do not play an important role in the institutional segment, in contrast to the evidence from the retail segment reported in Table 4. Furthermore, the coefficients of all other family-level variables—that is, Family Excess Return, Family Variance Excess Return, and Family Turnover—are insignificant in most specifications, providing further evidence that within-family spillovers in the institutional segment are not as relevant, as in the retail segment of the market.

Interestingly, Tables 2–4 show that the OLS and IV estimates are almost identical in the sample of retail funds, whereas Tables 5–7 show that the IV estimates tend to be lower (although, sometimes, not statistically different) than the OLS estimates in the sample of institutional funds. This "difference in difference" seems intuitive, as the economic effects that should generate a downward bias in the OLS coefficients—that is, families launching new funds to disguise the poor performance of existing funds or to maximize the chances of having a fund with superior returns—should be less relevant in the institutional segment, as the investors are, on average, more sophisticated than retail investors.

□ **Robustness checks and alternative hypotheses.** We now present robustness checks using an alternative set of instruments and discuss in detail two leading alternative hypotheses, presenting arguments against them.

Alternative instruments. We now verify the robustness of the results to several potential concerns about the validity of the instruments. As highlighted, the validity of the instruments relies on two main assumptions: (i) each fund family represents a small fraction of Employment Financial Sector in the cross-section of U.S. counties; and (ii) the exclusion restriction that, for example, Employment Financial Sector in Los Angeles or Boston does not *directly* affect the demand and market share of American Funds or Fidelity, whose headquarters are in Los Angeles and Boston, respectively. Although in principle, we could subtract each family's employees from Employment Financial Sector to directly avoid any reverse causality concern, CRSP does not, unfortunately, report this information, and we are not aware of another public data set that does. Moreover, this alternative procedure would not directly address concerns about the plausibility of our exclusion restriction. Furthermore, another potential concern about our instruments is that families' location is endogenous. For example, Employment Financial Sector could affect the number of family funds—as this article posits—and other unobserved dimensions of a family's "quality" of offerings, such as financial advice. If this is the case, the estimates of demand spillovers may be biased.

To address all these concerns, in the spirit of the empirical literature on differentiated products following Berry, Levinsohn, and Pakes (1995), we use supply-side variables of other families as instruments for each family's endogenous measure of varieties. Specifically, for each family i,

TABLE 7 Demand Spillovers: Category Regressions, Institutional Segment

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Log(Family Funds)	.1993		−.0177		−.3148		−.5099		.6529	
	(.0717)		(.2551)		(.4459)		(.4268)		(.5881)	
Log(Family Categories)		.1020		−.2752		−.3402		−.5285		.4375
		(.0950)		(.3215)		(.4610)		(.4549)		(.9822)
Log(Category-Family Fees)	−.8895	−.9340	−1.4793	−1.5673	−4.5194	−4.2032	−6.2816	−5.6920	−1.5826	−1.6001
	(.1508)	(.1510)	(.4454)	(.4477)	(1.7139)	(1.3311)	(1.8303)	(1.4888)	(.7525)	(1.0976)
Category-Family Return	.9984	1.0243	1.3569	1.3761	.6256	.6351	1.0448	1.0656		
	(.1468)	(.1470)	(.2309)	(.2332)	(.1323)	(.1219)	(.2388)	(.2372)		
Category-Family Turnover	−.0745	−.0687	−.0527	−.0464	.0441	.0304	.0890	.0654		
	(.0393)	(.0384)	(.0373)	(.0361)	(.0642)	(.0510)	(.0665)	(.0542)		
Category-Family Age	.1285	.1292	.1238	.1238	.0898	.0922	.0933	.0959	.1191	.1129
	(.0163)	(.0167)	(.0211)	(.0213)	(.0272)	(.0266)	(.0276)	(.0271)	(.0904)	(.1413)
Family Excess Return	1.1647	1.2695	−.5278	−.6100	−.8515	−.8302	−1.8038	−1.7185	.1494	1.3338
	(.6996)	(.7139)	(.5941)	(.5952)	(.6864)	(.6540)	(.8612)	(.8214)	(1.8640)	(1.7066)
Family Variance Excess Return	−2.7600	−3.2766	−.9032	−1.0242	1.5460	1.4454	2.7747	2.5603	−.1875	−.6705
	(1.6836)	(1.6910)	(1.3842)	(1.3162)	(1.6390)	(1.5791)	(1.7132)	(1.6431)	(9.9345)	(9.7873)
Family Turnover	.0223	.0362	−.0097	.0134	.0095	.0025	.1081	.0866	−1.0749	−.8771
	(.1272)	(.1312)	(.1511)	(.1461)	(.1991)	(.1872)	(.2048)	(.1928)	(.5240)	(.5302)
Year Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes	N/A	N/A	Yes	Yes
Category Fixed Effects	No	No	No	No	Yes	Yes	N/A	N/A	N/A	N/A
Category-Year Fixed Effects	No	No	No	No	No	No	Yes	Yes	N/A	N/A
No. Observations	12,061	12,061	12,061	12,061	12,061	12,061	12,061	12,061	263	263
No. Category-Family Pairs	1,971	1,971	1,971	1,971	1,971	1,971	1,971	1,971	38	38

Note: This table reports the results of OLS and IV regressions that investigate whether a category of institutional funds offered by a family that offers more funds or categories than its rivals in the institutional segment has a larger market share. Specifications (1) and (2) refer to OLS regressions, and specifications (3)–(10) refer to IV regressions. Specifications (1)–(8) employ the sample of all institutional category-family pairs. Specifications (9) and (10) employ the category of institutional S&P 500 Index funds. Each observation corresponds to a category-family-year triple. The dependent variable is Log(Cat-Fam Market Share), the log of a family-category's market share, in all specifications. The independent variables are defined in Section 5. In specifications (3)–(10), the endogenous variables Family Funds, Family Categories, and Category-Family Fees are instrumented using the instruments described in the text. All specifications also include year fixed effects; specifications (5) and (6) further include Lipper categories fixed effects; specifications (7) and (8) include fixed effects for each Lipper category-year pair. The data refer to the years from 1999 to 2007. Standard errors in parentheses are clustered at the fund family level.

TABLE 8 Alternative Instruments: Family Regressions

	(1)	(2)	(3)	(4)
Log(Family Funds)	1.5098		1.2186	
	(.1939)		(.5326)	
Log(Family Categories)		2.0809		1.0190
		(.2889)		(.6178)
Log(Family Fees)	−2.982	−2.6904	−3.4199	−3.5357
	(.5895)	(.6672)	(1.3596)	(1.3459)
Family Age	.0337	.0258	.0298	.0125
	(.0093)	(.0091)	(.0256)	(.0224)
Family Excess Return	.9060	.9725	1.5424	1.4645
	(.3126)	(.3402)	(.5509)	(.5387)
Family Variance Excess Return	.2614	.2772	.5687	.2540
	(.3068)	(.3178)	(1.3881)	(1.4170)
Family Turnover	.0186	.0196	−.0652	−.0476
	(.0405)	(.0430)	(.1330)	(.1368)
Year Fixed Effects	Yes	Yes	Yes	Yes
No. Observations	3,753	3,753	1,574	1,574
No. Families	528	528	231	231

Note: This table reports the results of Instrumental Variable regressions that investigate whether families that offer more funds or categories have a proportionally larger aggregate market share. Specifications (1) and (2) refer to the retail segment; specifications (3) and (4) refer to the institutional segment. Each observation corresponds to a family-year pair. The dependent variable is Log(Family Market Share), the log of each family's aggregate market share in a given segment. The independent variables are defined in Section 5. The endogenous variables Family Funds, Family Categories, and Family Fees are instrumented using the instruments described in the text. All regressions further include year fixed effects. The data refer to the years from 1999 to 2007. Standard errors in parentheses are clustered at the fund family level.

we compute the average Employment Financial Sector, Total Establishments, and Distance from NYC of all families whose headquarters are located in a *different* county from family i, and we use those as instruments for Family Funds and Family Categories of family i. In oligopoly markets, these are valid instruments because rival families' cost shifters affect equilibrium best responses and, thus, families' equilibrium characteristics, such as product varieties. In addition, by construction, these instruments do not suffer from reverse causality and do not directly affect families' demands and market shares. Moreover, as the empirical literature on differentiated products since Berry, Levinsohn, and Pakes (1995) advocates, they are arguably uncorrelated with each family's unobservable characteristics.

Table 8 reports the second stage of the family-level regressions using these alternative instruments. Columns (1) and (2) refer to retail funds, and columns (3) and (4) to institutional funds. The table shows that the estimates of the coefficients are almost identical to those reported in columns (3) and (4) of Tables 2 and 5, respectively, indicating that our results are robust to potential concerns about the validity of the instruments.

We have also performed the fund-level and category-level regressions using this alternative set of instruments. The results (omitted) confirm that, on average, the market share of a fund or of a family's category of funds is larger if the number of funds (categories) offered by its family is larger in the retail segment, but not in the institutional segment. (Details are available upon request.) These additional checks further corroborate the robustness of our results on the importance of demand spillovers in the retail segment, but not in the institutional segment.

Economies of scale and scope. Khorana and Servaes (1999) suggest that economies of scale and scope are important in the mutual fund industry. These economies may induce larger families to launch more new funds, and they may also constitute a barrier to entry for new families, thus potentially explaining some of the empirical patterns documented. Similarly, in an influential article, Schmalensee (1978) argues that incumbent firms may choose to offer multiple products in

order to "fill the product space," thus crowding out additional entrants. (Bonanno, 1987, analyzes Schmalensee's argument, showing that incumbents deter entry through product *specification*, rather than through product *proliferation*.)

Although economies of scale and scope are key features that shape competitive outcomes in the mutual fund industry, we reiterate that our empirical model identifies investors' *demand*, because the instruments used in our empirical model exploit exogenous supply-side shifters of the number of funds and number of categories offered by each family. In addition, our analysis reveals striking differences between the retail and institutional segments of the market. All supply-side factors—including economies of scale and scope—should not differ between retail and institutional funds. Moreover, our results stand when we restrict our analysis to (almost) homogeneous products such as S&P 500 Index funds. Furthermore, several families offer funds in both the retail and the institutional segments of the market, presumably sharing any costs savings due to lower research, product development, and administrative costs. Thus, it is not clear why the two segments of the market exhibit such stark differences. But our argument rests on demand spillovers, and we highlighted in Section 3 why these spillovers differ between the retail and institutional segments.

Advertising and research expenditures. Sutton's (1991, 1998) seminal work focuses on two main types of endogenous sunk costs: advertising and research outlays. Thus, a natural question is whether these costs can account for the empirical patterns documented. In particular, Park (2008) documents that advertising expenditure has increased over time in the mutual fund industry (in particular, for no-load funds), and this increase may have fostered concentration. However, it is unlikely that this alternative hypothesis can explain *all* our empirical findings, for several reasons. First, we wish to emphasize that our empirical model is designed to precisely control for spurious correlations due to unobserved factors, including advertising and research expenditures. In particular, the instruments that we employ in the empirical analysis exploit exogenous variations in the number of funds and number of categories offered. Second, many families offer funds in both the retail and the institutional segments of the market and, presumably, the effects of advertising and research (in particular) are not confined to a single segment of the market. Thus, it is not immediately obvious why market conduct and market structure respond differently to the same input. Third, and perhaps most important, Gallaher, Kaniel, and Starks (2009) investigate patterns of advertising in the mutual fund industry and find that families with funds in more objective classes advertise *less* than families with fewer objectives. Hence, if advertising were the key determinant of market structure and concentration in the retail segment, their finding would imply that families with funds in more objective classes should have a *larger* market share, in sharp contrast to the results of our analysis. Thus, we conclude that advertising and research expenditures cannot explain our empirical findings.

6. Conclusions

■ This article investigates, both theoretically and empirically, the role of demand for firms' product varieties and demand spillovers in determining market conduct and market structure in the mutual fund industry. The model adapts Sutton's (1991) endogenous sunk costs theory, highlighting that the magnitude of spillovers determines the industry equilibrium. If demand for firms' product portfolios and, thus, demand spillovers are strong, competition induces families to offer a large number of funds rather than low fees. As a result, the industry remains concentrated even in large markets. Instead, if demand for firms' product portfolios and spillovers are weak, the number of funds that each family offers is limited, fees are lower, and the industry is more fragmented in larger markets.

Aggregate empirical patterns reveal striking differences between the retail and institutional segments of the market: in the retail segment, families offer more funds; funds have higher fees; the market is more concentrated; and the total number of funds is larger than in the institutional segment. These patterns are exactly the outcomes predicted by the theoretical model if spillovers

are stronger in the retail than in the institutional segment. Indeed, the empirical analysis provides strong evidence that these spillovers are stronger in the retail segment of the market.

The ideas of this article are potentially useful in understanding several market outcomes in all industries in which consumers prefer to purchase from a single supplier (banking, commercial aircraft, supermarkets, etc.). For example, an interesting question, left for future research, is what are the determinants and the effects of mergers in such markets.

Appendix

First-stage regressions. Table A1 reports the results of the first-stage regressions of the endogenous variables Family Funds—that is, the total number of funds offered by a family—and Family Categories—that is, the total number of categories/investment objectives in which a family offers at least one fund—offered by family i in year t in the respective segment (retail or institutional) on the following instruments: Employment Financial Sector, the total employment in the financial sector in the same county in which the family has its headquarters; Total Establishments, the number of establishments in the same county in which the family has its headquarters; Distance from NYC, the distance of the family headquarters' address from New York City; Average Excess Return Other Families, the average Family Excess Return of all other families; and Average Variance Excess Return Other Families, the average Family Variance Excess Return of all other families. Specifications (1) and (3) refer to retail funds, and specifications (2) and (4) refer to institutional funds.

The signs of the coefficients of the instruments are largely as expected. In particular, on average, a larger employment in the financial sector, relative to the total employment, in the same county in which the family has its headquarters corresponds to a larger number of funds or to a larger number of categories offered by the family. This is true in both the retail and institutional segments. Moreover, the instruments are jointly significant: the F tests are equal to 24.06, 14.07, 23.91, and 13.22, respectively, in specifications (1)–(4).

To appreciate the magnitude implied by the coefficients, we construct the fitted number of funds from the coefficients of column (1). These fitted values imply, for example, that the average number of funds offered by a family whose

TABLE A1 The Relationship between Family Product Variety and the Instruments

	Log(Family Funds)		Log(Family Categories)	
	(1)	(2)	(3)	(4)
Family Age	−.0301	−.0336	−.0182	−.0224
	(.0017)	(.0042)	(.0013)	(.0031)
Family Excess Return	−10.5670	−6.6371	−6.7710	−7.9637
	(2.3411)	(5.5696)	(1.8144)	(4.5459)
Family Variance Excess Return	−9.6742	−71.726	−8.3555	−57.9082
	(5.1218)	(20.0828)	(3.8464)	(16.3562)
Family Turnover	−.0148	.0223	−.0182	.0143
	(.0145)	(.0394)	(.0013)	(.0318)
Log(Employment Financial Sector/100,000)	1.0205	.9400	.7604	.7216
	(.1118)	(.1499)	(.0814)	(.1181)
Log(Total Number of Establishments/100,000)	−1.2996	−1.0196	−1.0166	−.8258
	(.1659)	(.2308)	(.1220)	(.1772)
Average Excess Return Other Families	−4521.2	−1273.3	−2920.5	−1496.7
	(982.2)	(984.8)	(763.9)	(806.0)
Average Variance Excess Return Other Families	−3604.8	−11751.7	−3131.1	−9477.9
	(2084.3)	(3408.7)	(1559.3)	(2778.2)
Distance from NYC	.0026	.0037	.0025	.0028
	(.0016)	(.0021)	(.0012)	(.0017)
Year Fixed Effects	Yes	Yes	Yes	Yes
Industry Segment	Retail	Institutional	Retail	Institutional
No. Observations	3,753	1,574	3,753	1,574
No. Families	528	231	528	231

Note: This table reports the results of first-stage regressions. The dependent variable is the log of the number of retail funds offered by a family in specification (1); the log of the number of institutional funds offered by a family in specification (2); the log of the number of Lipper categories in which the fund family offers at least one retail fund in specification (3); and the log of the number of Lipper categories in which the fund family offers at least one institutional fund in specification (4). The independent variables are defined in the text. All regressions further include year fixed effects. The data refer to the years from 1999 to 2007. Robust standard errors in are parentheses.

headquarters are in New York, is twice as large as a family's whose headquarters are in Austin, TX: 33 funds versus 16.6 funds.

References

Agnew, J., Balduzzi, P., and Sundén, A. "Portfolio Choice and Trading in a Large 401(k) Plan." *American Economic Review,* Vol. 93 (2003), pp. 193–205.

Berry, S. "Estimating Discrete-Choice Models of Product Differentiation." *RAND Journal of Economics,* Vol. 25 (1994), pp. 242–262.

—— and Waldfogel, J. "Product Quality and Market Size." *Journal of Industrial Economics,* Vol. 58 (2010), pp. 1–31.

——, Levinsohn, J., and Pakes, A. "Automobile Prices in Market Equilibrium." *Econometrica,* Vol. 63 (1995), pp. 841–890.

Bonanno, G. "Location Choice, Product Proliferation and Entry Deterrence." *Review of Economic Studies,* Vol. 54 (1987), pp. 37–45.

Borenstein, S. "Hubs and High Fares: Dominance and Market Power in the U.S. Airline Industry." *RAND Journal of Economics,* Vol. 20 (1989), pp. 344–365.

——. "The Dominant-Firm Advantage in Multiproduct Industries: Evidence from the U. S. Airlines." *Quarterly Journal of Economics,* Vol. 106 (1991), pp. 1237–1266.

Choi, J., Laibson, D., and Madrian, B. "Plan Design and 401(k) Savings Outcomes." *National Tax Journal,* Vol. 57 (2004), pp. 275–298.

Christoffersen, S. and Sarkissian, S. "City Size and Fund Performance." *Journal of Financial Economics,* Vol. 92 (2009), pp. 252–275.

Cohen, L. and Schmidt, B. "Attracting Flows by Attracting Big Clients." *Journal of Finance,* Vol. 64 (2009), pp. 2125–2151.

Draganska, M. and Jain, D. "Product-Line Length as a Competitive Tool." *Journal of Economics and Management Strategy,* Vol. 14 (2005), pp. 1–28.

Economides, N. "Desirability of Compatibility in the Absence of Network Externalities." *American Economic Review,* Vol. 79 (1989), pp. 1165–1181.

Ellickson, P. "Quality Competition in Retailing: A Structural Analysis." *International Journal of Industrial Organization,* Vol. 24 (2006), pp. 521–540.

Elton, E., Gruber, M., and Blake, C. "The Adequacy of Investment Choices Offered by 401K Plans." *Journal of Public Economics,* Vol. 90 (2006), pp. 1299–1314.

Employee Benefit Research Institute. "Total Individual Account Retirement Plan Assets, by Demographics, 2004." *EBRI Notes*, Vol. 29, No. 3 (March 2008).

Fujita, M., Krugman, P., and Venables, A. *The Spatial Economy*. Cambridge, Mass.: MIT Press, 1999.

Gallaher, S., Kaniel, R., and Starks, L. "Advertising and Mutual Funds: From Families to Individual Funds." Mimeo, University of Texas, Austin, 2009.

Gil-Bazo, J. and Ruiz-Verdú, P. "The Relation between Price and Performance in the Mutual Fund Industry." *Journal of Finance,* Vol. 64 (2009), pp. 2153–2183.

Gowrisankaran, G. and Krainer, J. "Entry and Pricing in a Differentiated Products Industry: Evidence from the ATM Market." *RAND Journal of Economics,* Vol. 42 (2011), pp. 1–22.

Gremillion, L. *Mutual Fund Industry Handbook*. Hoboken, NJ: John Wiley and Sons, 2005.

Holmes, T. and Stevens, J. "Geographic Concentration and Establishment Scale." *Review of Economics and Statistics,* Vol. 84 (2002), pp. 682–690.

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

Huberman, G. and Jiang, W. "Offering vs. Choice by 401(k) Plan Participants: Equity Exposure and Number of Funds." *Journal of Finance,* Vol. 61 (2006), pp. 763–801.

Investment Company Institute. *Investment Company Fact Book.* Washington, DC: Investment Company Institute, 2008.

Ippolito, R. "Consumer Reaction to Measures of Poor Quality: Evidence from the Mutual Fund Industry." *Journal of Law and Economics,* Vol. 35 (1992), pp. 45–70.

Ivković, Z. "Spillovers in Mutual Fund Families: Is Blood Thicker Than Water?" Mimeo, University of Illinois, 2004.

Kadiyali, V., Vilcassim, N., and Chintagunta, P. "Product Line Extensions and Competitive Market Interactions: An Empirical Analysis." *Journal of Econometrics,* Vol. 89 (1998), pp. 339–363.

Katz, M. and Shapiro, C. "Network Externalities, Competition, and Compatibility." *The American Economic Review,* Vol. 75 (1985), pp. 424–440.

—— and ——. "Systems Competition and Network Effects." *Journal of Economic Perspectives,* Vol. 8 (1994), pp. 93–115.

Khorana, A. and Servaes, H. "The Determinants of Mutual Funds Starts." *Review of Financial Studies,* Vol. 12 (1999), pp. 1043–1074.

—— AND ——. "Competition and Conflicts of Interests in the U.S. Mutual Fund Industry." Mimeo, London Business School, 2007.

KLEMPERER, P. "Competition When Consumers Have Switching Costs: An Overview with Applications to Industrial Organization, Macroeconomics, and International Trade." *Review of Economic Studies,* Vol. 62 (1995), pp. 515–539.

KLEMPERER, P. AND PADILLA, A. "Do Firms' Product Lines Include Too Many Varieties?" *RAND Journal of Economics,* Vol. 28 (1997), pp. 472–488.

KNITTEL, C. AND STANGO, V. "Incompatibility, Product Attributes and Consumer Welfare: Evidence from ATMs." *B.E. Journal of Economic Analysis and Policy,* Vol. 8 (2008), Article 1.

—— AND ——. "How Does Incompatibility Affect Prices. Evidence from ATMs." *Journal of Industrial Economics,* Vol. 57 (2009), pp. 557–582.

LANCASTER, K. *Variety, Equity, and Efficiency*. New York: Columbia University Press, 1979.

LEDERMAN, M. "Are Frequent Flyer Programs a Cause of the 'Hub Premium'?" *Journal of Economics and Management Strategy,* Vol. 17 (2008), pp. 35–66.

MARSHALL, A. *Principles of Economics*, 8th edition. London: Macmillan, 1920.

MASSA, M. "Why So Many Mutual Funds? Mutual Funds, Market Segmentation and Financial Performance." Mimeo, INSEAD, 1997.

——. "How Do Family Strategies Affect Fund Performance? When Performance-Maximization is not the Only Game in Town." *Journal of Financial Economics,* Vol. 67 (2003), pp. 249–304.

MATUTES, C. AND REGIBEAU, P. "Mix and Match": Product Compatibility without Network Externalities." *RAND Journal of Economics,* Vol. 19 (1988), pp. 221–234.

MCFADDEN, D. "Conditional Logit Analysis of Qualitative Choice Behavior." In P. Zarembka, ed., *Frontiers in Econometrics*. New York: Academic Press, 1974.

NANDA, V., WANG, Z., AND ZHENG, L. "Family Values and the Star Phenomenon: Strategies of Mutual Fund Families." *Review of Financial Studies,* Vol. 17 (2004), pp. 667–698.

——, ——, AND ——. "The ABCs of Mutual Funds: On the Introduction of Multiple Share Classes." *Journal of Financial Intermediation,* Vol. 18 (2009), pp. 329–361.

PARK, M. "Advertising and Market Share Dynamics." Mimeo, University of California-Berkeley, 2008.

POLLET, J. AND WILSON, M. "How Does Size Affect Mutual Fund Behavior?" *Journal of Finance,* Vol. 63 (2008), pp. 2941–2969.

SCHMALENSEE, R. "Entry Deterrence in the Ready-to-Eat Breakfast Cereal Industry." *Bell Journal of Economics,* Vol. 9 (1978), pp. 305–327.

SHAKED, A. AND SUTTON, J. "Product Differentiation and Industrial Structure." *Journal of Industrial Economics*, Vol. 36, No. 2 (Dec., 1987), pp. 131–14.

—— AND —— "Multiproduct Firms and Market Structure." *RAND Journal of Economics,* Vol. 21 (1990), pp. 305–327.

SIALM, C. AND STARKS, L. "Mutual Fund Tax Clienteles." Mimeo, University of Texas Austin, 2009.

SIRRI, E., AND TUFANO, P. "Competition and Change in the Mutual Fund Industry." In S.L. Hayes, ed. , *Financial Services: Perspectives and Challenges*. Boston: Harvard Business School Press, 1993.

SIRRI, E. AND TUFANO, P. "Costly Search and Mutual Fund Flows." *Journal of Finance,* Vol. 53 (1998), pp. 1589–1622.

STARKS, L. AND YATES, M. "Reputation and Mutual Fund Choice." Mimeo, University of Texas, Austin, 2008.

SUTTON, J. *Sunk Costs and Market Structure*. Cambridge, Mass.: MIT Press, 1991.

SUTTON, J. *Technology and Market Structure*. Cambridge, Mass.: MIT Press, 1998.

WAHAL, S., AND WANG, A. "Competition among Mutual Funds." *Journal of Financial Economics,* Vol. 99 (2011), pp. 40–59.