Advertising and Market Share Dynamics

Minjung Park*

December 27, 2008

Preliminary and Incomplete; Do not quote without author's permission

Abstract

This paper examines the role of advertising in the evolution of market structure in the US mutual fund industry. First, I present empirical evidence that mass advertising in this industry creates endogenous sunk costs, resulting in patterns consistent with Sutton's (1991) predictions. In particular, I contrast evolutionary patterns of advertising spending and market structure between two segments of the industry, one in which mass advertising is effective (no-load segment) and the other in which mass advertising is much less so (load segment), due to differences in their distribution methods. I then estimate a dynamic model of advertising using the two-step estimator proposed by Bajari, Benkard, and Levin (2007). No-load firms and load firms face different advertising elasticities of demand because of the way their funds are marketed to consumers, and it leads to different optimal advertising choices and market structure dynamics in the two segments. I recover the structural parameters of interest, such as the advertising elasticity of demand in each segment, goodwill accumulation equation, entry costs, and sell-off values. Using the model estimates, I simulate market share dynamics when firms follow different advertising strategies. Thus, this paper provides insights on how strategic advertising decisions influence market structure.

1 Introduction

We know that sunk (and fixed) costs of incumbents could act as a barrier to entry. Many forms of advertising, e.g., mass advertising on TV or print, involve sunk costs, and it is well established that

*University of Minnesota. This is based on the third chapter of my Ph.D. dissertation. I thank Tim Bresnahan, Liran Einav, Jon Levin, Pat Bajari, and Kyoo il Kim for helpful conversations. Correspondence: mpark@umn.edu.

incumbents often use advertising as a strategic tool to maintain their dominance. As Sutton convincingly argued in his influential work (1991), in a growing market that can increasingly accommodate a larger number of firms, incumbents may escalate their advertising expenditures to keep a concentrated market structure.

Then, why do firms in some growing industries choose to employ this "escalation in advertising" strategy to maintain their dominance, while firms in other growing industries don't? Such a difference in firm behavior arises because firms face different incentives depending on various industry features, such as the shape of advertising cost function, how relevant "fixed cost" advertising is for the industry, consumers' responsiveness to advertising, etc. For example, if we have two growing industries that differ in consumers' responsiveness to mass advertising but are otherwise similar, by comparing mass advertising behavior of incumbents and resulting market share dynamics between the two industries, we can infer the relationship among consumers' response to advertising, firms' optimal advertising choice, and market structure.

This research attempts to investigate such a relationship exploiting a nice feature in the US mutual fund industry. The mutual fund industry can be naturally divided into two segments: load segment and no-load segment. Load firms sell their mutual funds via brokers, while no-load firms sell their funds directly to consumers (investors). Because investors in the load segment heavily rely on brokers for their choice of mutual funds, they are not very responsive to mass advertising. Instead, a load firm hires wholesalers who market its products to brokers on an individual basis, and the wholesalers are paid "incentives" depending on how many funds they convince brokers to sell to consumers. Hence, a large fraction of marketing costs for load firms fall on the variable costs part. On the other hand, mass advertising is the dominant form of marketing in the no-load segment and could be quite effective in raising consumers' willingness to pay or awareness since consumers choose which brands to buy from without any professional help. Hence, a large portion of marketing costs fall on the fixed costs part for no-load firms. I exploit this natural divide of the mutual fund industry into two segments—one that is more like an exogenous sunk costs market (load segment)[1] and the other which is closer to being an endogenous sunk costs market (no-load segment). From 1985 to 2004, market demand in both segments grew significantly, mainly due to exogenous factors such as a long bull market in the 1990s and an explosive growth of defined contribution pension plans. The research question posed in this paper is then as follows. Given the demand growth and the difference in consumer responsiveness to mass

[1] We have "exogenous" sunk costs if the effectiveness of fixed outlays in raising product quality or reducing unit costs is arbitrarily low (Sutton, 2006).

advertising in these two segments, how do the evolutions of advertising patterns and market structure differ between the two segments?

More specifically, Sutton's model of endogenous sunk costs yields the following testable predictions. In the no-load segment where consumers are responsive to mass advertising, firms who were dominant at the beginning of the sample period would keep increasing their ad spending as the market grows. This escalation in endogenous fixed costs by dominant firms would deter entry into the top-tier by other firms such as fringe firms or new entrants. As a result, the market will not fragment despite the large increase in market size. On the other hand, in the load segment where consumers are much less responsive to mass advertising, we would expect a much smaller increase in dominant firms' advertising spending over time compared to the no-load segment. As a result, the load segment is likely to get less concentrated with the increase in market size.

Those predictions are borne out in the data. From 1985 to 2004, market size, measured by assets under management, increased from \$0.44 trillion to \$3.37 trillion and from \$0.31 trillion to \$3.13 trillion for the load segment and no-load segment, respectively.[2] With this increase in market size, the load segment became more fragmented over time, with the 3-firm concentration ratio (5-firm concentration ratio) declining from 33 to 25.15 (44.76 to 31.78). In contrast, the no-load segment became *more* concentrated despite the significant increase in market size. The 3-firm concentration ratio (5-firm concentration ratio) for the no-load segment *increased* from 37.83 to 52.81 (from 48.41 to 60.39). Unlike the contrasting evolutions for the dominant firms of the two segments, the fringe of the two segments, which do not rely on mass advertising for their survivals, evolved similarly over time: A large number of small firms entered the fringes of both segments as the markets grew larger.

The advertising patterns of dominant firms are also consistent with the predictions. The largest firms in the no-load segment spend more on mass advertising than the largest firms in the load segment, and the discrepancy in their ad spending widened significantly over time, with the largest no-load firms increasing their ad spending much more than the largest load firms did. This indicates that the no-load segment experienced an escalation in fixed investments, in this case brand enhancement via advertising. Moreover, this escalation in fixed investments is limited to dominant firms; fringe firms of either segment did not experience such an escalation in advertising spending. I also find that there is a much higher correlation between big companies, measured by asset size, and big ad spenders in the no-load segment than in the load segment. This suggests that mass advertising is a more important determinant of

[2] In 1998 dollars

dominance in the no-load segment.

After carefully documenting these data patterns, I estimate a model of advertising. My model is based on the dynamic model of Ericson and Pakes (1995) and I estimate the model using the two-step estimator proposed by Bajari, Benkard, and Levin (2007). I also borrow modeling insights from the Oblivious Equilibrium of Weintraub, Benkard, and Van Roy (2008a) and its extensions (WBV, 2008b) to allow for a two-tiered structure in the market—dominant tier and fringe. This extension is important since the mutual fund industry has more than a few hundred firms in total, among which only a few are dominant players and the rest are very small fringe firms. Another important feature is that market size grows in my application.[3] My dynamic programming problem is still stationary in the sense that the value function or policy function does not explicitly depend on time t, once I model how the functions depend on market size. An implication of a growing market is that I need to be able to project the optimal advertising decisions for market sizes unobserved in the data, and my simple solution is to parameterize the policy functions.[4]

In the model firms make advertising, pricing, entry, and exit decisions, and the industry structure is determined as a result of these decisions as well as optimal purchasing decisions by consumers. A firm's current advertising spending affects not only its current product demand but also future demand via the impact of advertising on the firm's goodwill stock. No-load firms and load firms face different advertising elasticities of demand because of the way their funds are marketed to consumers, and this leads to different optimal advertising choices and market structure dynamics in the two segments. I recover the structural parameters of interest, such as the advertising elasticity of demand in each segment, the goodwill accumulation equation, marginal costs of production, entry costs, and sell-off values. The model allows me to infer how much of one-shot ad spending is required for a new entrant to have the same level of goodwill stock as the largest incumbents in the industry. This figure measures endogenously created sunk costs of entry into the top-tier of the market. The model also allows me to simulate market share dynamics when firms use different advertising strategies. My simulation shows that if no-load firms were not allowed to engage in any mass advertising, market structure in the no-load segment would have been much more fragmented than as we observe it today. Hence, the results provide insights on how strategic advertising decisions influence market structure.

[3] Market size, however, will eventually stop growing and the value function is bounded.

[4] The need to parameterize the policy function in the first-step of BBL estimation arises in any application where some states are not realized in the data or the state vector does not exhibit a stationary Markov transition (e.g., growing network in Ryan and Tucker (2008)). Hence, the need for parameterization is not unique to my application. I am being explicit about it because a growing market is an important feature of my data.

This paper makes contributions to three strands of literature. The first is the scarce empirical literature on advertising in a dynamic setting. A few papers in this literature are Doraszelski and Markovich (2007), Dubé, Hitsch, and Manchanda (2005), and Doganoglu and Klapper (2006). There are a few papers that study advertising in the context of the mutual fund industry, although the setting is not dynamic (Cronqvist, 2006; Gallaher, Kaniel, and Starks, 2006; Jain and Wu, 2000).

The second literature my paper contributes to is the literature on sunk costs and market structure (Shaked and Sutton, 1983, 1987; Sutton, 1991). Papers following these pioneering works have empirically shown the impact of endogenous sunk costs on market structure. Ellickson (2004, 2007) argues that the supermarket industry remains concentrated even in large markets because in larger markets incumbent firms incur greater investments in quality through ever more advanced distribution centers. Berry and Waldfogel (2006) find that the newspaper industry, where investments in quality mainly come through fixed costs, such as more or better staffs, remains concentrated in large markets, while the restaurant industry, for which quality improvements predominantly fall on variable costs, such as better materials, becomes fragmented as the market grows. Bresnahan and Greenstein (1999) apply Sutton's theory to explain why the number of platforms is limited in the computer industry.

Finally, this paper contributes to the fast developing literature on estimation of dynamic games. Ericson and Pakes (1995) proposed a model of firm dynamics in an oligopoly setting, and many papers since then have extended and/or estimated EP-style dynamic games (Benkard, 2004; Doraszelski and Satterthwaite, 2007; Gowrisankaran and Town, 1997; Pakes and McGuire, 1994, 2001; Weintraub et al., 2008). The biggest challenge to using EP-style models has been computation, and recent papers have developed estimation methods that do not require repeated computation of equilibria, hence are computationally much more feasible (Aguirregabiria and Mira, 2007; Bajari, Benkard, and Levin, 2007; Pesendorfer and Schmidt-Dengler, 2004). I employ estimation methods of Bajari et al. (2007) in this paper. Papers that use the two-step estimation methods include Beresteanu and Ellickson (2007), Collard-Wexler (2007), Macieira (2006), Ryan (2006), Ryan and Tucker (2007), and Sweeting (2007) among others. Weintraub, Benkard, and Van Roy (2008a, 2008b) develop a model of industry dynamics with many firms in which firms do not need to follow competitors' individual state variables and call the equilibrium concept Oblivious Equilibrium (OE). I use a concept similar to OE to model the fringe in the mutual fund industry. Empirical papers that use the OE framework include Xu (2008).

The rest of this paper proceeds as follows. Sections 2 and 3 describe the mutual fund industry and my data sets. Section 4 presents time-series patterns of advertising behavior and industry structure. Section

5 discusses a dynamic model of advertising and empirical specifications. Section 6 presents estimation results. Section 7 concludes the paper.

2 Industry

2.1 Load vs. No-Load and Mass Advertising

Load firms and no-load firms differ in their distribution methods. Load firms sell funds through brokers. They market their products to brokers, who then choose which mutual funds to recommend to their customers. The customers pay loads (commissions) to brokers for their advice. Investors who would prefer professional advice on their investment choices tend to choose load funds. In order to market their funds to brokers, load firms hire wholesalers who travel across the country to pitch the firms' products to brokers, build relationship, and provide auxiliary support to the brokers such as seminars. The wholesalers typically receive base salary and incentive pay. This marketing style is very similar to pharmaceutical companies' hiring of detail reps to market their products to physicians. Because load firms need to appeal to brokers rather than consumers directly, mass advertising is not their main marketing strategy. Load firms do engage in some mass advertising, because of their need to reach those who invest through defined contribution plans (I discuss defined contribution plans in detail below) and also in order to build brand recognition among investors who then might ask their brokers to invest in those funds for them. However, as Gallaher et al. (2006) put it, "Load funds rely more on the brokers and dealers, rather than advertising, to reach their investors."

On the contrary, no-load firms sell their mutual funds directly to individual investors. Investors who would rather save on commissions and not receive professional advice tend to choose no-load funds. For no-load firms, mass advertising is a very effective way of increasing brand power and awareness of their products among a large audience. Mass advertising might also provide useful information to consumers, such as past performance as well as overall fee structure.[5] No-load firms advertise on newspapers, magazines, and TV, and papers have shown that advertising is helpful in attracting capital from investors (Cronqvist, 2006; Gallaher et al., 2006; Jain and Wu, 2000; Reuter and Zitzewitz, 2006). In other words,

[5] In this paper, I do not try to distinguish between "informative" and "prestige" effects of advertising. I believe both effects are present in this market, although brand effects seem to be a much stronger driver of advertising, as I discuss below. The lack of consumer-level data prevents me from distinguishing between the two effects, as was proposed by Ackerberg (2001).

no-load firms can increase demand for their products via heavy advertising which involves an increase in fixed costs. Certainly not all investors are influenced by advertising, and there are firms who cater to these brand-insensitive customers (and do not advertise). Consequently, the no-load segment can be considered as consisting of two tiers, one tier dominated by large firms who advertise heavily to attract brand-sensitive consumers and the other tier filled with many fringe firms who do not advertise and appeal to brand-insensitive consumers.

Although the load segment and no-load segment are quite distinct from each other, the distinction is becoming less clear-cut, as fund complexes start to use multiple channels of distribution. For instance, both load and no-load firms sell funds through defined contribution pension plans, which is neither a direct channel nor a brokerage channel. No-load funds might be sold through intermediaries such as fee-based financial planners, fund supermarkets, or wrap programs. This might necessitate marketing to these financial planners by no-load firms. Conversely, appealing to individuals who invest in mutual funds via defined contribution plans might necessitate mass advertising for load firms. The growing tendency to utilize multiple channels of distribution blurs the distinction between load firms and no-load firms, and as a reflection of this, a few no-load firms have launched funds that carry loads to tap into the demand of load funds. For the most part, however, there is a clear enough distinction, and I categorize each firm as either a load firm or no-load firm based on its portfolio of funds. I discuss the criterion I use for the categorization in Section 3.

One might wonder why some investors are willing to pay commissions to brokers to buy load funds when they could presumably buy other no-load funds that invest in similar asset classes with similar performance. The commissions could be as high as 5% of the total investments, which is not a trivial amount especially considering the fact that the typical annual management fees that fund companies charge for their actively managed funds are about 1.5%. Bergstresser, Chalmers, and Tufano (2004) try to answer this question by examining possible benefits that brokers might deliver to investors: superior asset allocation, access to better performing funds or funds that are harder to find, attenuation of investor bias, etc. Interestingly they do not find any evidence of tangible benefits of broker advice that could justify the commission investors pay to brokers. Given the rapid expansion of the pool of individuals who invest in mutual funds, the lack of investment savvy for novice investors might provide a partial explanation. In this paper, I do not try to answer this question. I just take as given the presence of different types of consumers and examine how their response to advertising differs and how that difference might affect the optimal advertising decisions of firms in each segment of the market.

Advertising decisions are made at the firm level, rather than at the fund level by individual fund managers. Hence, my unit of analysis will be firm, not fund. Gallaher et al. (2006) note that "In conversation with mutual fund family executives, they indicated that the intent of the advertising is often not the particular fund advertised, but the fund family itself." Similarly, Gremillion (2005) notes "Fidelity's advertising emphasized the Fidelity brand, much as consumer product firms used branding to sell shoes or shop. While many fund complexes today do the same thing, it was Fidelity under Ned Johnson that led them to it."

Mass advertising is only part of all marketing expenditures that firms incur. There are no precise statistics on this, but according to a 1999 survey by Financial Research Corporation and PricewaterhouseCoopers, about 12% of marketing expenditures (excluding commissions paid to brokers) by 24 fund companies, some load firms and others not, were spent on mass advertising, while the rest mostly went to staff costs, including salary and incentives for wholesalers. I do not have data on marketing expenditures other than mass advertising, so I focus on mass advertising decision in this paper. Because my main interest is how the presence of endogenous sunk costs, or lack thereof, influences market structure, and mass advertising is a big part of such endogenous sunk costs, focusing on mass advertising seems to be an appropriate choice.

2.2 What Changed in the Past 20 Years

In this subsection, I discuss major changes that occurred in this industry in the past 20 years. To understand why the industry structure evolved the way it did and why the evolution was different in the two segments, one needs to know how the environment in which firms operate, including demand and technology of distribution, has changed. I discuss two biggest changes that might have had an impact on the industry structure. One is an explosive growth of demand, especially from defined contribution pension plans, and the other is a new distribution system called fund supermarket that revolutionized fund distribution, particularly for fringe firms. There are other, less prominent factors that might have affected the market structure differently in the two segments, such as different merger rates. I delay discussion of these minor factors to Section 4 where I present descriptive results.

2.2.1 Retirement Savings

Tax-advantaged retirement investments have been the main driving force behind the growth of the mutual fund industry for the last 20 years. In 1985, there were almost no retirement assets held in mutual funds. At the end of 2003, retirement assets held in mutual funds totaled about $2.7 trillion or 36 percent of total mutual fund assets. The $2.7 trillion assets divided almost evenly into two major categories: individual retirement accounts (IRAs) and employer-sponsored defined contribution plans (Gremillion, 2005). For the traditional IRA, individuals make tax-deferred contributions to investment vehicles they choose, up to a limit set by law. Individuals can choose from all investment vehicles available on the market, such as individual securities, mutual funds, bank deposits, etc., and employers do not play any role. It is up to individuals to choose what investments to make and set up accounts with financial institutions to execute the investment decisions.[6] Therefore, assets that flow into mutual funds through IRAs are not different from other mutual fund assets in the sense that the investors could choose from any mutual fund.

Employer-sponsored defined contribution plans, among which 401(k) is the most famous, are different. For defined contribution plans, an employer sponsors a retirement plan that is administered by a certain record keeper and has a fixed menu of investments. Employees then choose only from the investment options available in the plan their employer sponsored. These plans typically include only a limited number of investment options which encompass the sponsoring employer's own stock, bank deposits, and mutual funds. Elton et al. (2006) found that the median number of investment options in their sample of pension plans is eight. Moreover, the mutual funds offered by a plan tend to come from only a few mutual fund companies, if not one. According to Gremillion (2005), "At one time, the large mutual fund companies offering record keeping service could insist that only their mutual funds be included in the plan's investment choices. Until 1995, for example, Vanguard required plans for which it performed record-keeping to choose mutual funds only from the Vanguard family. Competition in recent years has forced record keepers to accommodate a wide range of investment choices by plans, including funds from multiple complexes. For example, in 2000 the Delaware Group performed record keeping for the PricewaterhouseCoopers 401(k) plan, but the plan's investment options included funds managed by American Express, Neuberger Berman, Lazard, Northern Trust, and Brinson, in addition to Delaware." Even these days, however, the number of fund complexes included in a typical pension plan is very low.

[6] There are other kinds of IRA where employers set up accounts for their employees. But they are very small compared to the traditional IRA.

What this implies for industry structure is significant. Since defined contribution plans tend to include funds from the largest fund complexes only, and employees who invest through the plans must choose from the fixed menus, mutual fund companies that were large when DC plans got popular were at a significant advantage in attracting this fast growing pool of retirement assets. In other words, there was a demand-side change that might have affected market structure and it could confound the impact of advertising on the market structure of the two segments. In Section 4, I examine how assets invested through defined contribution plans might have differentially affected market structure in the load and no-load segments.

2.2.2 Fund Supermarkets

There now exist distribution channels of mutual funds that did not exist or play any meaningful role 20 years ago, such as defined contribution plans, wrap programs, fee-based financial advisors, and fund supermarkets. Among them, the fund supermarket is the most significant development, next to defined contribution plans. Pozen, in his book (2002) on the mutual fund industry, notes "Today's popular version of the mutual fund supermarket was introduced by discount brokerage firm Charles Schwab in 1992 and has since transformed the way investors purchase and sell funds." Fund supermarkets are platforms, usually online these days, where investors can buy funds from various fund families. The popular version Pozen refers to is NTF (no transaction fee) supermarket. If a fund participates in an NTF supermarket, investors can transact in the fund without paying any trading fees. The fund supermarket charges participating fund companies an annual fee of 25 to 35 basis points for providing the platform and performing customer account services. The share of NTF supermarket distribution has gone from zero to 5% of the industry total from 1992 to 2002 (Reid and Rea, 2003). In 2002, the two largest NTF supermarkets, Schwab One-Source and Fidelity Funds Network, held more than $150 billion in assets.

Most funds sold through supermarkets are no-load funds, and one would expect that fund supermarkets might have affected the market structure of the no-load segment (but no direct impact on the load segment). Particularly, fund supermarkets made it easier for small entrants to penetrate the no-load market. By putting their funds on fund supermarkets, small no-load firms could reach consumers without heavy advertising for brand-name buildup. As Standard & Poor's Investment Industry Survey (1999) puts it, "the supermarket format actually levels the playing field by giving a small fund as much public visibility as a large one." However, since there is a large portion of assets that do not flow via fund supermarkets, firms who depend mainly on supermarkets for distribution and do not advertise would not be able to break into the top tier of dominant firms. In other words, fund supermarkets might have made easier entry into the fringe part of the no-load segment, but likely did not have much impact on concentration among top-tier no-load firms.

3 Data

I use three sources of data for my analysis: CRSP data, Ad $ Summary data, and Fed's Flow of Funds. I describe each data set in turn. First, comprehensive data on U.S. mutual funds are available from the Center for Research in Security Prices (CRSP). The CRSP data set includes information on all open-end mutual funds that have ever existed including: the amount of assets invested by the fund, the identity of the management company running the fund, the fund's investment objective, the fund's monthly returns, and the structure of the fund's fees (loads, expense ratios). The data are at the fund level, and I can aggregate up to the firm level using the management company information of each fund.

The identifiers CRSP assigns management companies, however, are not necessarily unique since CRSP reuses the identifiers of extinct management companies. Thus, I constructed a unique management company identifier to be able to track companies over time. CRSP does not provide the identity of the management company for each fund in years prior to 1992, but I obtained such information for earlier years by matching fund names with another data set from Thomson Financial. From the CRSP data set, I can obtain statistics on industry structure such as the distribution of market shares, the numbers of entrants and exiters, the number of acquisitions, etc. I also observe characteristics of each firm, including its portfolio of funds and their characteristics.

As I briefly mentioned in the previous section, I categorize each firm as a load firm or a no-load firm. Since the data set is at the fund level and records whether a fund carries loads, I can compute the proportion of funds that carry loads among all funds offered by a given fund family. In doing so, I ignore money market funds since money market funds almost never carry loads even when they are offered by a load company. This is due to the short-term nature of investments in money market funds. People invest in money market funds as a way to park their money temporarily while looking for good investment possibilities. Hence charging loads on money market funds would be very unpalatable to investors, and as a result money market funds are mostly no-load in the data, regardless of firms who offer them. I also ignore institutional funds and retirement plans in the computation, since these types of funds typically do not carry loads regardless of firms who offer them. Load firms tend to waive their loads for institutional investors or DC plan investors.

Another issue I had to deal with is that sometimes no-load funds impose early redemption penalties (which are different from back-end loads) to discourage short-term trading, and that these early redemption penalties are not distinguishable from regular loads in the data. Using the fact that these early redemption penalties tend to be much smaller than typical loads (redemption fees tend to be less than 2% of assets) and considering the entire portfolio of a given firm to see the pattern of loads among its funds, I determine whether a fund is a load fund, or a no-load fund simply carrying early redemption fees. Once having dealt with these issues, I obtain the proportion of load funds for each firm. If the proportion is more than 0.9 (35% of data), I categorize the firm as a load firm. If the proportion is equal to zero (55% of data), I categorize the firm as a no-load firm. Indeterminate cases (about 10% of data) arise primarily because of the failure to perfectly distinguish between early redemption fees and loads or to sort out money market / institutional / DC plan funds. Carefully checking each firm usually allows me to determine whether a firm is a load firm or not with a reasonable degree of certainty. There are about a dozen firms who seem to use both the load and no-load distribution channels extensively. For my analysis, I assume that these firms are no-load firms.

My second data set contains information on fund companies' mass advertising spending from Ad $ Summary collected by Competitive Media Reporting (CMR).[7] CMR measures space for all ads that appear in major media and multiply it by appropriate rates to obtain ad expenditures by firm and brand in various media. In 1998, the CMR data set covered 10 major media—consumer magazines, Sunday magazines, newspapers, outdoor, network television, spot television, syndicated television, cable television, network radio, and national spot radio. The coverage of the data set is quite comprehensive. For instance, CMR monitored over 220 consumer magazines, 255 newspapers editions, 37 cable television networks, and so on. However, CMR does not know the exact rate each advertiser pays, and relies on going rates to compute ad expenditures by firm. I collected the data from Ad $ Summary books for 1985 - 2001. From 2002 on, the data were available in an electronic format, so I use them.

Often a mutual fund company is part of a bigger financial institution (e.g., Morgan Stanley offers its own mutual funds) and these financial institutions do general advertising of the companies as a whole, in addition to specifically promoting their mutual funds. CMR reports figures for this type of general promotion separately from figures for mutual fund ads. It does not make sense to include these gen-

[7] Until 1992, the data were collected by National Leading Advisers.

eral promotion figures in my analysis since the mutual fund business might be only a small part of the whole company, in which case the benefit of the general promotion mainly accrues to other products the firm offers. On the other hand, entirely excluding the figures might also cause a problem since general advertising of Morgan Stanley might make consumers more willing to buy funds from Morgan Stanley. The best solution would be to figure out the contribution of the mutual fund business for each financial institution and assign appropriately adjusted ad expenditures. However, collecting this information is very difficult. Hence, I will exclude ads that generally promote entire financial institutions and just focus on ads specifically promoting their mutual funds.

The last data set is on potential market size for the mutual fund industry. My definition of the potential market size is financial assets held by domestic financial sector, and I obtain this measure from annual publications of the Fed's Flow of Funds.

4 Empirical Facts

4.1 Industry Structure

In this subsection I examine the market structure of the mutual fund industry. The realized demand for mutual funds increased dramatically for both the load segment and the no-load segment during the sample period of 1985 through 2004. Table 1[8] shows that over the sample period the assets under management increased by $2.9 trillion and $2.8 trillion for the load segment and no-load segment, respectively. Over the course of 19 years, the markets continued to increase without interruption, except for the recession period following the dot com bubble burst in 2000.

With this drastic increase in market size, one would expect market fragmentation to a certain degree. In the load segment, the market does get fragmented somewhat. In Table 2, we see that dominant firms became less dominant, and the market accordingly became more fragmented over time in the load segment. The 3-firm concentration ratio (C_3) and 5-firm concentration ratio (C_5) decreased from 33 to 25.15 and from 44.76 to 31.78 respectively over time. On the contrary, the no-load segment became more concentrated despite the drastic increase in market size. The 3-firm concentration ratio and 5-firm concentration ratio for the no-load segment *increased* from 37.83 to 52.81 and from 48.41 to 60.39,

[8] All figures in my analysis are in 1998 dollars.

respectively. Although not shown, the identities of the top 5 players in each segment remained very stable throughout the sample period, with firms who were dominant at the beginning of the sample period still being dominant at the end.

Unlike the contrasting evolutions for the dominant firms of the two segments, the fringe of the two segments, which do not rely on mass advertising for their survivals, evolved similarly over time: A large number of small firms entered the fringes of both segments as the markets grew larger. Table 3 shows such a pattern. We also see that although the number of firms occupying the fringe increased for both segments, the increase was much greater for the no-load segment. This is consistent with the conjecture that fund supermarkets made easier entry into the fringe part of the no-load segment. Particularly, we observe a very high rate of net entry into the fringe of the no-load segment between 1992 (year in which the first NTF fund supermarket appeared) and 1999 (before the onset of the dot com bubble bust), while the number of fringe firms in the load segment did not change much during the same period. Before 1992, the number of fringe firms increased similarly in the two segments.

As well, Table 3 shows the same pattern we saw in Table 2 from a different angle. I count the number of largest firms in each segment who, combined together, serve 50% of the market. In the load segment 6 largest firms served 50% of the market in 1985, but by 2004 12 firms divided 50% of the market. In other words, the individual shares of dominant players shrank over time. In the no-load segment, the number of largest firms whose combined market shares reach 50% decreased from 6 in 1985 to 3 in 2004. This shows that dominant players became more dominant over time in the no-load segment. Another pattern emerging from Table 3 is that the relative size ratio between dominant firms and fringe firms shrinks over time in the load segment, but the ratio is getting larger in the no-load segment. This again shows that the market gets fragmented with an increase in market size in the load segment, but not in the no-load segment.

4.2 Advertising Patterns

In this subsection I discuss advertising patterns. The patterns would suggest that an escalation in fixed investments, in this case via mass advertising, in one segment and its lack in the other might explain the different evolutions of market structure in those segments. Table 4 reports the number of advertisers in each segment. The table shows that the load and no-load segments tend to have similar numbers of advertisers, which steadily increased over time. On the other hand, the no-load segment experienced a much larger increase in the number of non-advertisers compared to the load segment. This suggests that there has been a high rate of entry by fringe no-load firms who do not rely on advertising to reach consumers. The particularly large increase in the number of non-advertisers in the no-load segment between 1993 and 1997 might be due to fund supermarkets which made easier entry into the fringe part of the no-load segment. The existence of many firms who do not advertise at all suggests that the fringe firms cater to brand-insensitive consumers.

Table 5 reports ad-sales ratios for the top 5 players and the rest in each segment, respectively. The typical definition of sales does not seem appropriate for this industry since firms' revenue comes from annual fees charged to both existing fund shares and new shares sold. Hence, in computing ad-sales ratios, I define sales as the total assets under management times the expense ratios (expense ratios are annual fees investors pay for management of their money, expressed in percentage of assets). We see from Table 5 that ad-sales ratios are much higher for no-load firms than for load firms and also that the largest players tend to have higher ad-sales ratios than the other smaller players in each segment.[9]

Table 6 shows how advertising spending has changed over time. The first and third columns are the average advertising spending by the five biggest firms (measured by the amount of assets under management) within the load segment and no-load segment, respectively. The second and fourth columns are the average advertising spending by the rest of the firms in each segment, including those who do not advertise. Two patterns stand out in the table. The largest no-load firms spend more on mass advertising than the largest load firms, and the discrepancy in their ad spending widened significantly over time, with the largest no-load firms increasing their ad spending much more than the largest load firms did. The discrepancy increased from \$0.89 million in 1985 to \$9.9 million in 2004. This is consistent with the no-load segment experiencing an escalation in fixed costs investments, in this case brand enhancement via advertising: Dominant no-load firms strategically choose to fortify their status as dominant ones by incurring large fixed costs to raise their products' perceived quality and/or consumer awareness. This escalation in fixed costs investments, however, is limited to dominant firms. Small firms in the no-load segment, who presumably serve relatively brand-insensitive consumers, spend less on mass advertising than small firms in the load segment throughout the sample period, and the increase in ad spending by these small firms was much smaller than the increase experienced by large players.

Table 7 shows that there is a tight relationship between big companies (by asset size) and big ad spenders in the no-load segment, while the relationship is much less tight in the load segment. For

[9] The absolute levels of ad-sales ratios might not seem high compared to other "high advertising" industries. However, note that the definition of sales I use is somewhat different from the typical definition of sales. If I define ad-sales ratios using the revenue from only new fund shares sold, the ratio would be much higher.

instance, in 2004, all of the five biggest no-load firms are among the top ten ad spenders within the no-load segment, while only two of the five biggest load firms are among the top ten ad spenders within the load segment. This suggests that mass advertising is a much more important determinant of dominance in the no-load segment than in the load segment.

So far I have shown that the divergence in the industry structures of the two segments might be explained by differences in their mass advertising behavior. The argument, à la Sutton, is that as market size increases over time, incumbents would have an incentive to increase their investments in quality if the investments occur mainly through fixed costs and are effective at raising consumers' willingness to pay for their products. In the no-load segment, a large portion of investors respond to mass advertising, and mass advertising expenditures are fixed costs, so incumbent dominant firms increase their ad spending with an increase in market size. As a result, the market remains concentrated despite the increase in market size. In the load segment, people are much less influenced by mass advertising because they rely on brokers to make a choice. Thus, incumbent dominant firms cannot deter other firms from entering the top tier by mass advertising, and an increase in market size leads to a fragmented market structure.

There are other potential explanations for the observed evolutions in the industry structure. One is that the drastic growth of defined contribution plan assets might have induced a higher concentration in the no-load segment. To check this possibility, I compute C_3 and C_5 for each segment excluding DC plan assets. The idea is that if the no-load segment is concentrated because DC plan assets are concentrated in a few top no-load firms, I would observe a less concentrated no-load market once I exclude DC plan assets. To do this, I need to know how much of DC plan assets each fund company manages. For 2003, I know the top ten leading fund companies for management of DC plans and how much of DC plan assets each of them managed. I can also compute 2003 market size without DC plans by subtracting DC plan assets from total assets in each segment.[10] Using these, I can check how the market structure roughly would have looked without DC plan assets. The result shows that C_3 and C_5 remain almost unchanged when I exclude DC plan assets in their computation, suggesting that DC plan assets cannot be an explanation for the different market structures of the two segments and especially the high concentration rate in the no-load segment.

Another possibility is that mergers, which are fairly common in this industry as discussed in Park (2008), might have occurred more frequently among large no-load firms than large load firms, again

[10] The distribution of DC plan assets among the top 10 DC plan fund managers shows that about 70% of these top 10 firms' DC plan assets are managed by no-load firms (4 firms) and 30% are managed by load firms (6 firms). I assume that the same 70-30 division holds for the overall DC plan assets.

leading to a more concentrated structure in the no-load segment. I check this possibility by comparing the frequency and size of acquisitions by the top 3 firms in the no-load segment to those of acquisitions by the top 3 firms in the load segment. Among these firms, only Franklin Investment (load company) has engaged in an acquisition of a firm with more than $5 billion assets. Also, the amount of assets acquired by the top 3 load firms is larger than the amount of assets acquired by the top 3 no-load firms, suggesting that M&A cannot explain the concentrated market structure of the no-load segment.

5 Model

5.1 Setup[11]

My model is based on the dynamic oligopoly framework of Ericson and Pakes (1995). I also borrow insights from Weintraub, Benkard, and Van Roy (2008a, 2008b). I consider a discrete time model with infinite horizon and index times by $t \in \{0, 1, 2, \ldots\}$. Each firm that operates in the industry is indexed by a unique integer i. I let $\mathcal{I}_t$ denote the set of indices available in the industry at time t.

Dynamics in the model arises from the carryover effects of advertising as well as entry and exit. Current advertising affects not only current profits, but also future profits through its impact on the stock of goodwill. Ideally, I would like dynamics to arise from state dependence in demand as well because consumers do not shuffle their mutual fund portfolios frequently due to tax consequences, desire to avoid additional commissions, etc. Incorporating demand dynamics, however, is not an easy task for my application since it is challenging to distinguish between state dependence and unobserved heterogeneity, especially without individual-level data. Hence, I assume static demand in this paper: Mutual fund investors are assumed to make their mutual fund choices every period anew.

There are two tiers in the industry: "top tier" and "fringe." We can think of fringe players as mainly catering to brand-insensitive consumers, as evidenced by lower ad-sales ratios of smaller firms in Table 5. Since much of advertising is done by the largest players, in my model only dominant firms have the option to advertise (Below I will describe how I empirically define dominant players). Fringe firms do not make active advertising or pricing decisions. There are, however, active margins of action for fringe firms: entry and exit. I assume that entry and exit occur only for fringe firms, and not for dominant firms.

[11] Much of discussion in this subsection is based on work in progress by Kim and Park (2008). In that paper, we study a model of industry dynamics with a growing market and two-tiered structure. The model I consider in this subsection is a simplified version of the model studied in that paper.

In other words, a potential entrant can enter the market only as a fringe firm, and only a fringe firm can exit the market. The asymmetric treatment of dominant and fringe firms—dominant firms make pricing and advertising decisions only; fringe firms make entry and exit decisions only—is reasonably justified by the data, as I rarely observe exits or entry by firms that I empirically define as dominant players. Another key distinction between dominant firms and fringe firms is that dominant firms' state variables are monitored by every firm in the market, while only some summary statistics of fringe firms' state variables, rather than the whole distribution, are monitored by firms in the market (WBV, 2008b). Using this approximation in a two-tiered framework is necessary to make the model tractable since there are more than a few hundred mutual fund companies in a given year. In the model, a fringe firm remains fringe forever and a dominant firm remains dominant forever. Although the assumption of fixed tier is less desirable than endogenous switching in tiers, I note that the model still allows endogenous determination of industry structure through other channels. Industry concentration will endogenously change over time due to changes in the individual market shares of dominant firms via their advertising and pricing decisions as well as entry and exit decisions by fringe firms.

I collect in x_{it} all firm-specific state variables of period t. x_{it} contains firm i's tier ($\tau_i = 0$ if fringe firm, $\tau_i = 1$ if dominant firm), the firm's goodwill stock (G_{it}), the quality of its product (μ_{it}), and its type ($L_i = 0$ if no-load firm, $L_i = 1$ if load firm). I allow some elements of x_{it} to be continuous. For example, the level of goodwill stock can take any non-negative value. Using $|\cdot|$ to denote the size of a vector or a set, $|x_{it}|$ represents the dimension of x_{it} vector. Let $\mathcal{I}_t^D$ denote the set of indices available for dominant firms in the industry at time t and $\mathcal{I}_t^F$ the set of indices of fringe firms. Let N_t^D denote the number of dominant incumbents and N_t^F the number of fringe incumbents in period t. I define $N_t = N_t^D + N_t^F$. All random variables in the model are defined on a probability space $(\Omega, \mathcal{F}, \mathcal{P})$ equipped with a filtration $\{\mathcal{F}_t : t \geq 0\}$. All the random variables subscripted by t are $\mathcal{F}_t$-measurable. The industry is described by the following model primitives:

$$\begin{gathered}\{\pi(x_{it}, \mathbf{x}_t^D, \mathbf{x}_t^F, z_t), p_{\mathbf{x}^D}(\mathbf{x}_{t+1}^D|\mathbf{x}_t^D, \mathbf{x}_t^F, z_t), p_{\mathbf{x}^F}(\mathbf{x}_{t+1}^F|\mathbf{x}_t^D, \mathbf{x}_t^F, z_t), p_z(z_{t+1}|\mathbf{x}_t^D, \mathbf{x}_t^F, z_t)\} \\ (\mathbf{x}_t^D, \mathbf{x}_t^F, z_\mathbf{t}) \in \Omega^{|x_{it}|^{N_t^D}} \times \Omega^{|x_{it}|^{N_t^F}} \times \Omega\end{gathered} \tag{1}$$

The primitives in (1) are common knowledge among all actual and potential participants in the industry. The term $\mathbf{x}_t^D$ is the collection of dominant firms' state vectors and the term $\mathbf{x}_t^F$ is the collection of fringe firms' state vectors. The term z_t is the state vector which is common to all firms. In my application, I will interpret z_t as market size and accordingly treat z_t as a scalar, $z_t \in R_+$.

$\pi(x_{it}, \mathbf{x}_t^D, \mathbf{x}_t^F, z_t)$ represents firm i's expected per-period profit when its state vector is x_{it}, the state vectors for dominant firms and fringe firms (including its own) are $\mathbf{x}_t^D$ and $\mathbf{x}_t^F$, respectively, and market size is z_t. Later I will write down more fundamental constructs such as consumer utility function and product market competition from which π is derived, but for now I just work with π with an implicit understanding that π is consistent with some consumer utility function and product market competition. $p_{\mathbf{x}^D}(\mathbf{x}_{t+1}^D|\mathbf{x}_t^D, \mathbf{x}_t^F, z_t)$ is the Markov transition of $\mathbf{x}_t^D$ and $p_{\mathbf{x}^F}(\mathbf{x}_{t+1}^F|\mathbf{x}_t^D, \mathbf{x}_t^F, z_t)$ is the Markov transition of $\mathbf{x}_t^F$. $\mathbf{x}_t$ and z_t could influence the value of $\mathbf{x}_{t+1}$ indirectly through the actions taken by firms in period t, or directly, or both. Hence, it is understood that $p_{\mathbf{x}^D}(\cdot)$ and $p_{\mathbf{x}^F}(\cdot)$ already take into account the impact of firm actions on the evolution of the state vector. In particular, the optimal entry and exit decisions of fringe firms are incorporated in $p_{\mathbf{x}^F}(\cdot)$ so that N_{t+1}^F might differ from N_t^F. $p_z(z_{t+1}|\mathbf{x}_t^D, \mathbf{x}_t^F, z_t)$ is the Markov transition of z_t. In my empirical application, I will assume that the evolution of market size is not influenced by firm-specific state vectors or firm actions and simply follows an exogenous Markov process, i.e., $p_z(z_{t+1}|\mathbf{x}_t^D, \mathbf{x}_t^F, z_t) = p_z(z_{t+1}|z_t)$.

In each period, a fringe incumbent i privately learns its sell-off value ϕ_{it}, drawn from a distribution $F_\phi(\phi)$. Fringe firm i will exit the market if and only if the sell-off value exceeds the continuation payoff from remaining in the industry. I allow load and no-load firms to draw their sell-off values from different distributions, $F_\phi^L(\phi)$ for load firms and $F_\phi^{NL}(\phi)$ for no-load firms. Dominant firms do not make exit decisions and always stay in the market.[12]

Each dominant firm decides how much to advertise in each period after it learns a piece of private information about its advertising costs, ν_{it}^A. Advertising increases the level of goodwill stock a firm owns, which raises awareness or perceived quality of its product among consumers. The option to advertise is available to dominant firms only. When dominant firm i spends A_{it} on advertising in period t, i's level of goodwill stock in the next period, G_{it+1}, is determined by

$$G_{it+1} = H(G_{it}, A_{it}) \tag{2}$$

where H denotes goodwill accumulation function.

In each period, there are many potential entrants. Each potential entrant has an assigned type— either a load firm or a no-load firm— and can enter the market only as the given type. If a firm decides to enter the market, it needs to pay sunk costs of κ_t. If the potential entrant's type is a load firm, its sunk cost is κ_t^L, which is the same for all other potential entrants of the same type. If the potential entrant's type is a no-load firm, its sunk cost is κ_t^{NL}, which is again the same for all other potential entrants of the same type. Let $F_\kappa^L(\kappa)$ and $F_\kappa^{NL}(\kappa)$ denote sunk cost distributions for the two types. A

[12] Since I rarely observe exits by dominant firms in the data, recovering the distribution of sell-off values for dominant firms would be difficult, if not impossible.

firm knows the realization of its own type's entry sunk cost κ before making an entry decision. Firms are allowed to enter only as a fringe firm. This assumption is justified because the level of advertising spending necessary to enter as a dominant firm is likely to be too high and risky to be made in one shot, especially if the entrant is unsure about its overall quality. If a firm enters the market in period t, it will start to earn profits from the following period on until its exit. Each potential entrant i observes its state vector for the next period x^e_{it+1}, drawn from a distribution $F_e(x^e)$, only after it makes its entry decision. I again allow different distributions for different types, $F^L_e(x^e)$ and $F^{NL}_e(x^e)$.

Each firm, whether it is an incumbent, potential entrant, fringe firm, or dominant firm, aims to maximize its expected net present value. In each period, the timeline of the events is as follows:

1. At the beginning of period t, all firms, both potential entrants and incumbents, observe the state vector $(\mathbf{x}^D_t, \mathbf{x}^F_t, z_t)$.

2. Each incumbent fringe firm i privately learns its sell-off value ϕ_{it} and then decides whether to exit or not. The exit decision is not revealed to other firms.

3. Potential entrants make entry decisions and the number of entrants is determined. The entry decision is not revealed to other firms. All entrants enter as a fringe firm.

4. Each incumbent dominant firm i privately learns its advertising cost shock, ν^A_{it}, and chooses the optimal level of advertising, A_{it}. Goodwill stocks are updated according to (2) and become known to all market participants.

5. Each incumbent dominant firm i learns its marginal cost shock, ν^P_{it}. These marginal cost shocks are known to other market participants as well. Incumbent firms compete in the spot market à la static Bertrand-Nash. Firms receive profits π, which are net of production costs as well as advertising costs, if any.

6. Exits occur and exiting firms receive their sell-off values plus the current spot market profits.

7. Entry occurs and new entrants pay an entry cost of κ. New entrants observe their state vectors for the next period x^e_{it+1}. The industry takes on a new state vector $(\mathbf{x}^D_{t+1}, \mathbf{x}^F_{t+1}, z_{t+1})$.

I assume that all the random variables—ϕ_{it} (sell-off values), ν^A_{it} (advertising cost shocks), ν^P_{it} (marginal cost shocks), and κ_t (entry sunk costs)—are iid for all t and i, and have finite expectations with well-defined distribution functions. The random variables are also assumed to be independent of one another.

$H(A_{it})$ is non-decreasing in A_{it} and $H(A_{it})$ takes positive values with strictly positive probabilities for all $A_{it} > 0$. $H(A_{it})$ is uniformly bounded above by $\overline{H} > 0$ and below by $\underline{H} < 0$ for all A_{it}, and the advertising spending A_{it} is also uniformly bounded by $\overline{A}$ for all i and t. The transitions generated by $H(A_{it})$ are assumed to be unique investment choice admissible. This assumption is a technical condition from Doraszelski and Satterthwaite (2007) to ensure that firms' advertising decision problem has a unique solution, generating a pure advertising strategy. I impose a boundedness condition such that z_t is uniformly bounded by $\overline{z}$ for all t. This condition ensures that the value function does not explode.

Equilibrium of this model is similar to a combination of Markov Perfect Equilibrium (MPE) and Oblivious Equilibrium (OE). In particular, I focus on equilibrium where each dominant firm's state vector is tracked by every firm in the market while only some *summary statistics* of fringe firms' state vectors (rather than the entire distribution) are monitored. A firm's strategy depends on the firm's state vector x_{it}, the entire distribution of dominant players' state vectors $\mathbf{x}_t^D$, and some simple statistics of fringe firms' state vectors, denoted by $\overline{\mathbf{x}}_t^F$, and market size z_t. Conditioning strategies on simple statistics instead of the entire distribution of fringe state vectors significantly reduces the dimensionality problem. In this paper, I will call these strategies *behavioral strategies* and the corresponding equilibrium *Behavioral Equilibrium* (BE). Since Behavioral Equilibrium is based on summary statistics of each period, it is different from OE which is based on the long-run average. BE concept is used in Krusell and Smith (1998), where they solve a stochastic growth model assuming that agents use strategies that are conditioned on some simple statistics of the entire wealth distribution in the economy. I use BE instead of OE as an equilibrium concept for my application since BE can more easily incorporate aggregate shocks or growing markets. A disadvantage of BE compared to OE is that there is no formal asymptotic result which shows that BE would be close to MPE as the level of market size goes to infinity (more discussion is provided in Kim and Park, 2008). For OE, Weintraub, Benkard, and Van Roy (2008a) prove such an asymptotic result. Hence, I resort to "behavioral" arguments to justify the use of BE: Since there are more than a few hundred fringe firms in each period and each of them is small, firms can make near-optimal decisions by conditioning their strategies on summary statistics of fringe firms' state vectors. An alternative interpretation of BE would be that it is econometricians' approximation: Firms play MPE, but it is difficult for econometricians to solve the equilibrium of a dynamic game with more than a few hundred firms. Therefore, we econometricians approximate the maximization problems using value functions which depend upon summary statistics of fringe state vectors.

I now define behavioral strategies. Given the assumption of privately known random sell-off values, this game always has equilibrium in a pure exit strategy, namely a cutoff exit strategy. The assumption of unique investment choice admissible transition $H(A_{it})$ ensures that firms' advertising strategies are pure strategies as well. I further assume that firms use symmetric pure strategies for their advertising and exit decisions. Let $\mathcal{M}$ denote the set of behavioral exit and advertising strategies taken by incumbents. An element $\sigma \in \mathcal{M}$ is defined by a pair $\sigma = (A, \chi)$, where A is an advertising strategy (null for fringe firms) and χ is an exit strategy (null for dominant firms). Given the assumptions of the model, there is a common behavioral advertising strategy among dominant firms $A(x_i, \mathbf{x}^D, \overline{\mathbf{x}}^F, z, \nu_i^A)$, and there is a common behavioral exit strategy among fringe firms $\chi(x_i, \mathbf{x}^D, \overline{\mathbf{x}}^F, z, \phi_i) \in \{0, 1\}$ where $\chi = 1$ indicates exit. The cutoff exit strategy states that a fringe incumbent firm $i \in \mathcal{I}_t^F$ exits at time t if and only if the sell-off value ϕ_{it} is greater than or equal to the continuation payoff which is a function of $(x_{it}, \mathbf{x}_t^D, \overline{\mathbf{x}}_t^F, z_t)$. This exit rule generates a random exit time of fringe firm i as $\varrho_i = \min\{t : \chi(x_{it}, \mathbf{x}_t^D, \overline{\mathbf{x}}_t^F, z_t, \phi_{it}) = 1\}$. Since dominant firms never exit in my model, for $i \in \mathcal{I}^D$ we have $\chi = 0$ for all t and $\varrho_i = \infty$.

Regarding entry, I assume that there are a large number of potential entrants who play a symmetric pure entry strategy. These firms can enter as a fringe firm only. Let $\lambda^L\left(\mathbf{x}^D, \overline{\mathbf{x}}^F, z\right)$ denote the expected number of entrants of load type at state $(\mathbf{x}^D, \overline{\mathbf{x}}^F, z)$ that results from the entry strategy. Let $\lambda^{NL}\left(\mathbf{x}^D, \overline{\mathbf{x}}^F, z\right)$ denote the expected number of entrants of no-load type at state $(\mathbf{x}^D, \overline{\mathbf{x}}^F, z)$ that results from the entry strategy. These aggregate entry rates $\lambda^L(\cdot)$ and $\lambda^{NL}(\cdot)$ will be endogenously determined, and our solution concept will require that they satisfy a free entry condition, as in Bresnahan and Reiss (1991). The set of behavioral entry rate functions is denoted by Λ. Note that all the behavioral strategies—advertising (A), exit (χ), and entry (λ)—are a function of $\overline{\mathbf{x}}^F$, summary statistics of the fringe, rather than $\mathbf{x}^F$.

With these objects in hand, we can define *behavioral value function.* If firm i follows strategy $\sigma' \in \mathcal{M}$, the other firms follow a common strategy $\sigma \in \mathcal{M}$, and the aggregate entry rate is $(\lambda^L, \lambda^{NL}) \in \Lambda$, we can define a behavioral value function as follows:

$$
\begin{aligned}
& V(x_i, \mathbf{x}^D, \overline{\mathbf{x}}^F, z | \sigma', \sigma, \lambda^L, \lambda^{NL}) \\
= E & \left[\sum_{k=t}^{\varrho_i} \beta^{k-t} \pi(x_{ik}, \mathbf{x}_k^D, \overline{\mathbf{x}}_k^F, z_k) + \beta^{\varrho_i - t} \phi_{i\varrho_i} \middle| x_{it} = x_i, \mathbf{x}_t^D = \mathbf{x}^D, \overline{\mathbf{x}}_t^F = \overline{\mathbf{x}}^F, z_t = z \right]
\end{aligned} \tag{3}
$$

where the expectation $E[\cdot]$ is taken with respect to the strategy of firm i, the strategy followed by its competitors, the entry rate function, and other random shocks in the economy. β is the discount factor. This value function should be interpreted as the expected net present value of a firm whose state vector is x_i and who follows behavioral strategy σ', under the assumption that the *true* state vectors of dominant firms and fringe firms are $\mathbf{x}^D$ and $\overline{\mathbf{x}}^F$. I will let $V(\cdot|\sigma, \sigma, \lambda^L, \lambda^{NL}) = V(\cdot|\sigma, \lambda^L, \lambda^{NL})$ if firm i also follows strategy σ.

We are now ready to define the solution concept: Behavioral Equilibrium consists of a strategy $\sigma \in \mathcal{M}$ and an entry rate function $(\lambda^L, \lambda^{NL}) \in \Lambda$ that satisfy the following conditions:

1. Firm strategies optimize a behavioral value function:
$$\begin{gathered} \sup_{\sigma' \in \mathcal{M}} V(x_i, \mathbf{x}^D, \overline{\mathbf{x}}^F, z|\sigma', \sigma, \lambda^L, \lambda^{NL}) = V(x_i, \mathbf{x}^D, \overline{\mathbf{x}}^F, z|\sigma, \lambda^L, \lambda^{NL}) \\ \forall (x_i, \mathbf{x}^D, \overline{\mathbf{x}}^F, z) \in R^{|x_i|} \times R^{|x_i|^{N^D}} \times R^{|\overline{\mathbf{x}}^F|} \times R_+ \end{gathered} \tag{4}$$

2. At each state, the expected numbers of entrants $(\lambda^L, \lambda^{NL})$ are such that either (a) the behavioral expected value of entry for entering firms is nonnegative and no additional firm could enter and earn nonnegative behavioral expected value, or (b) no firm could enter and earn nonnegative behavioral expected value: for $\forall (\mathbf{x}_t^D, \overline{\mathbf{x}}_t^F, z_t) \in R^{|x_i|^{N^D}} \times R^{|\overline{\mathbf{x}}^F|} \times R_+$,
$$\begin{gathered} \text{(a) } \beta E[V(x_{it+1}^e, \mathbf{x}_{t+1}^D, \overline{\mathbf{x}}_{t+1}^F, z_{t+1}|\sigma, \lambda^L, \lambda^{NL})|\mathbf{x}_t^D, \overline{\mathbf{x}}_t^F, z_t] - \kappa^j \geq 0 \text{ for } j = L, NL \\ \beta E[V(x_{it+1}^e, \mathbf{x}_{t+1}^D, \overline{\mathbf{x}}_{t+1}^F, z_{t+1}|\sigma, \lambda^L + 1, \lambda^{NL})|\mathbf{x}_t^D, \overline{\mathbf{x}}_t^F, z_t] - \kappa^L < 0 \\ \beta E[V(x_{it+1}^e, \mathbf{x}_{t+1}^D, \overline{\mathbf{x}}_{t+1}^F, z_{t+1}|\sigma, \lambda^L, \lambda^{NL} + 1)|\mathbf{x}_t^D, \overline{\mathbf{x}}_t^F, z_t] - \kappa^{NL} < 0 \\ \lambda^L > 0,\ \lambda^{NL} > 0 \\ \text{or (b) } \beta E[V(x_{it+1}^e, \mathbf{x}_{t+1}^D, \overline{\mathbf{x}}_{t+1}^F, z_{t+1}|\sigma, 1, \lambda^{NL})|\mathbf{x}_t^D, \overline{\mathbf{x}}_t^F, z_t] - \kappa^L < 0 \text{ if } \lambda^L = 0 \\ \beta E[V(x_{it+1}^e, \mathbf{x}_{t+1}^D, \overline{\mathbf{x}}_{t+1}^F, z_{t+1}|\sigma, \lambda^L, 1)|\mathbf{x}_t^D, \overline{\mathbf{x}}_t^F, z_t] - \kappa^{NL} < 0 \text{ if } \lambda^{NL} = 0 \end{gathered} \tag{5}$$

In (5), we implicitly have $\tau_i = 0$ in x_{it+1}^e inside the value function since firms can enter as a fringe firm only. It is straightforward to show that BE exists under mild technical conditions. With respect to uniqueness, we have no reason to believe that in general there is a unique BE, similarly with MPE and OE.

5.2 Empirical Specification

I estimate the model using the two-step estimator proposed by Bajari, Benkard, and Levin (2007). In this subsection, I discuss empirical specifications and estimation procedures. BBL proceeds in two steps. First, one recovers reduced-form policies (advertising, entry, and exit choices) as a function of state variables, and computes transition functions for the state variables. Consumer demand and static pricing game are also estimated in the first step. In the second step, given the recovered policy functions, transition probabilities, demand estimates, and marginal cost estimates, one finds values of structural parameters, such as distribution parameters for entry sunk costs, advertising costs, and sell-off values, which make the observed policy functions optimal. Forward simulation is used for the second step.

In my application, the set of dominant firms is assumed to consist of firms who have *ever* been one of the top 20 firms in the market during the sample period of 1989 through 2004.[13] For this group of firms, entry and exit occur rarely,[14] so I do not model entry or exit for dominant firms. This yields 30 dominant firms in total.[15] The rest will be treated as fringe firms, and I allow entry and exit on the fringe. As a result, industry concentration in the model will change over time due to changes in the individual market shares of dominant firms as well as entry and exit on the fringe. Since mutual funds are sold nationwide, there is only one geographic market, national, and each period defines a market.

Let N_t^D denote the number of dominant firms in the market in year t ($N_t^D = 30$ for $\forall t$) and N_t^F denote the number of fringe firms in year t. Each firm's state vector s_{it} consists of the following variables: its own tier ($\tau_i = 1$ for dominant firm, $\tau_i = 0$ for fringe), its own type ($L_i = 1$ for load firm, $L_i = 0$ for no-load firm),[16] its own goodwill stock (G_{it}; G_{it} is 0 for fringe firms since fringe firms do not advertise), its own quality (μ_{it}),[17] (L_i, G_{it}, μ_{it}) of other dominant firms (which I denote by $(L_i, G_{it}, \mu_{it})^D$), summary statistics of (L_i, G_{it}, μ_{it}) of other fringe firms, and market size (M_{it}) which is common to all firms. For summary statistics of fringe states, I assume that market participants track the number N_t^F and average quality $\overline{\mu}_t^F$ of fringe firms of each type. I denote the collection of individual state vectors s_{it} by $s_t \in S$.

The transitions of the state variables are assumed to be as follows. A given firm's tier and type (τ_i, L_i) are fixed over time. A firm's quality (μ_{it}) stochastically evolves according to an AR(1) process, $F_\mu(\mu_{it+1}|\mu_{it})$, independent of actions. The transition of quality μ_{it} depends on how various firm characteristics, such as returns, risks, the existence of star funds, etc., evolve. Obviously, to a certain degree

[13] I do not use years prior to 1989 in estimation of the dynamic game because, as explained below, I need ad spending for the previous few periods to construct goodwill stock for each period.

[14] Among those who have ever been one of the top 20 firms, there was one entry (Barclays), and three exits (all due to acquisitions by other large firms) during the sample period. I drop the entry case from my sample of dominant firms and instead include it in the fringe part of the market. For the firms that disappeared due to acquisitions, I treat those target firms as if they were part of the acquirers from the beginning so that I do not need to explicitly deal with mergers while still including those large firms in the top tier of the market.

[15] These 30 firms make up about 62-84% of total advertising spending.

[16] In reality, firms choose their types. In this paper, I do not model firms' choice of type, and just take firm type as something exogenously given.

[17] This quality measure summarizes the firm's overall characteristics (other than tier and goodwill stock) that might influence consumer demand, such as firm age, the number of offered products, etc. Below I describe how I back out each firm's μ_{it} from demand estimation. In the data, some fringe firms do engage in advertising. Since I do not allow fringe firms to accumulate goodwill in my model, their goodwill in the data will be captured by μ_{it}.

μ_{it} is influenced by firm actions such as hiring of good fund managers, but endogenizing these margins of actions is beyond the scope of this paper. By modeling μ_{it} to follow an AR(1) process, I allow persistence in firms' overall quality.

For market size M_{it}, I fit an AR(1) for first-differenced market size. I chose this specification because I was not able to reject the null hypothesis of unit root for AR(2) of the original series. This series, unfortunately, means that the market will grow without bound. To theoretically ensure that the value function is bounded above, I assume that there is an upper limit for the market size.[18] Practically, I choose a very large number for the cap of the market size in my forward simulation, although in all simulation runs I have done so far I never hit the bound.

Regarding goodwill stock, I assume that G_{it} deterministically evolves as a finite distributed lag of advertising.

$$G_{it} = \sum_{k=0}^{4} \delta^k A_{it-k} \tag{6}$$

The specification assumes that advertising in the previous four periods and the current period determine the goodwill stock of firm i. I chose a finite lag to avoid the initial condition problem. Alternatively, I could use an infinite lag and impose a distributional assumption on the initial goodwill stock, as in Dubé, Hitsch, and Manchanda (2005). The retention rate of advertising is captured by δ. Mainly for computational simplicity, I assume that goodwill stock evolves deterministically. I.e., there is no stochastic term in the goodwill accumulation function.

Potential entrants have an assigned type, load or no-load, and observe sunk costs of entry κ before making entry decisions. However, they observe the realization of other state variables only after they enter the market. Since they can enter the market only as a fringe firm, they face no uncertainty over τ_i and G_{it+1} (fringe firms do not advertise, so their goodwill stocks are always equal to zero). The only state variable they do not observe at the time of entry decision is quality μ. New entrant i observes its quality for the next period μ_{it+1} after entering the market, and I assume that μ_{it+1} for new entrants of each type are drawn from the empirical distribution of μ among fringe incumbent firms for the given type in period $t+1$.

Consumer demand is modeled using a logit model for differentiated products. Fund companies are multi-product firms (some firms offer more than 200 mutual funds), so I model consumer demand for fund companies rather than funds in order to simplify my analysis and make the unit of analysis consistent with

[18] A weaker condition is discussed in Kim and Park (2008). In this paper, I simply impose that there is an upper bound for market size.

the rest of the model. In other words, I treat each firm as a single product. Although an approximation, this captures the fact that many consumers concentrate their holdings in at most a few fund companies. To further simplify the analysis, I assume that each consumer chooses one fund company and that each consumer invests the same amount of money. A consumer who invests a large amount of money into mutual funds can be interpreted in my model as multiple investors each of whom invests the same amount. Since I use aggregate data to estimate demand, allowing for multiple discrete choices as in Hendel (1999) does not add much value. Each consumer chooses the fund complex that yields the highest utility for him. The utility of the consumer from choosing firm i depends on observed characteristics of the firm, unobserved (to econometricians) firm quality (ξ_{it}), and a logit error term which captures an idiosyncratic preference shock. Consumers can choose the outside good as well. I define the outside good to be financial investments other than mutual funds, and let choice 0 denote the outside good. The utility of individual h from choosing firm i in period t is given by

$$u_{hit} = (\beta_1 + \beta_2 L_i)\ln(G_{it}) - \alpha P_{it} + \gamma_1 L_i + \gamma_2 X_{it} + \xi_{it} + \epsilon_{hit} \tag{7}$$

One of the characteristics that enter the consumer utility function is firm i's level of goodwill stock. I allow demand sensitivity to goodwill stock to differ for load firms and no-load firms. If β_2 is less than zero, it suggests that no-load firms face higher advertising elasticity of demand than load firms. I take the log of goodwill stock to ensure that the marginal effect of advertising on consumer demand is decreasing.

There are two prices relevant for consumers' decisions, $P_{it} = (P_{1it}, P_{2it})$. P_{1it} is loads. These are non-zero only for load firms, obviously. Among load firms, investors would prefer those that charge lower loads, all else equal. A complication is that investors who buy load funds seek advice from brokers, who might prefer high-load funds. The loads are paid by consumers to brokers, rather than to fund companies, because they are fees for brokers' advice. To the extent that this agency relationship creates misaligned incentives, investors might be steered into funds with high loads. P_{2it} is expense ratios.[19] Investors, either investing in load funds or no-load funds, would prefer lower expense ratios, all else equal.

X_{it} includes, among other things, firm i's past performance, the number of fund offerings of the firm, the firm age, and the variety of fund objectives offered by the fund family. Since I as an econometrician do not observe ξ_{it} and prices are set reflecting this unobserved quality, there is a potential endogeneity problem. I deal with the problem by using IV after transforming market shares so that the unobserved

[19] To be precise, P_{2it} is the percentage of the total investment that fund investors pay for the mutual fund's operating expenses, including 12b-1 fees which are fees for distribution.

quality term appears linearly, as in Berry (1994). I use characteristics of other firms as instruments for prices. For the same reason, the level of goodwill stock, i.e., current and previous advertising expenditures, is also endogenous, and I need to instrument for it as well. Unfortunately, results are sensitive to which set of BLP instruments is used for advertising (results are not sensitive which set of BLP instruments is used for prices), so for now I do not instrument for advertising. The utility of the outside good is normalized to zero.

I estimate the above demand function (7) after plugging in the expression for goodwill stock (6), and this allows me to recover the parameter δ in the goodwill accumulation function as well as the parameters in the consumer utility function. The number of firm characteristics that enter the demand function is large. This helps us to estimate a realistic demand function but poses a serious problem for the overall dynamic model, because all the observed firm characteristics and unobserved firm quality that enter the demand function should be included in the state vector, rendering the dimension of the state vector too large.[20] To deal with this dimensionality problem while keeping demand estimation realistic, I construct a single-index quality measure μ_{it} for each firm using the demand estimates. Specifically, quality μ_{it} of firm i is $\mu_{it} = \gamma_1 L_i + \gamma_2 X_{it} + \xi_{it}$.[21] I then include these adjusted "mean utilities" μ_{it} in the state vector of the dynamic model. This approach allows me to convert the multi-dimensional firm characteristics into a single index of firm quality. The idea is similar to using inclusive value terms for nested logit models or multi-product firms (e.g., see Hendel and Nevo, 2006; Macieira, 2006). By including μ_{it} in the state vector of the dynamic game, I effectively allow serially correlated unobserved firm heterogeneity, exploiting the panel structure of the data.

I estimate the demand model assuming that ϵ_{hit} follows Type I extreme value distribution. The logit specification places strong restrictions on how unobservable product space increases with the number of products. To address this issue, I include $ln(N)$—where N is the number of available firms—as a regressor following Ackerberg and Rysman (2005).

For the period profit function, I assume constant marginal production costs and specify firm i's

[20] I can recover the unobserved firm qualities from demand estimation and include them in the state space treating them as if they are observed.

[21] When I estimate demand, I do not distinguish between dominant firms and fringe firms. After I do demand estimation, however, I need to construct μ_{it} for each firm and here I distinguish between dominant firms and fringe firms. Since I explicitly model advertising and pricing decisions of dominant firms, the quality measure μ_{it} for them should be net of price and goodwill effects. Hence, $\mu_{it} = \gamma_1 L_i + \gamma_2 X_{it} + \xi_{it}$ for dominant firms. For fringe firms, I do not model their pricing or advertising decisions. Therefore, the quality measure μ_{it} (which is assumed to evolve exogenously) for fringe firms should include price and goodwill effects that exist in the data: $\mu_{it} = (\beta_1 + \beta_2 L_i)\ln(G_{it}) - \alpha P_{it} + \gamma_1 L_i + \gamma_2 X_{it} + \xi_{it}$.

marginal production costs as a function of its type L_i. There is an iid shock to marginal costs which is normally distributed with mean 0 and variance $\sigma^2_{\nu^P}$.[22] Firms incur advertising costs as well. Because my advertising data are reported in dollar values, I do not need to estimate the per-unit cost of advertising. Instead, I allow for the possibility that there is some fixed costs firms need to incur to have positive amount of advertising. Moreover, there is an iid private shock to advertising costs which is normally distributed with mean 0 and variance $\sigma^2_{\nu^A}$. In the demand function, the impact of advertising on demand (hence profits) was assumed to be concave. To the extent that the logarithm transformation of goodwill stock in the demand function does not fully capture decreasing marginal returns to advertising, we would want to be flexible in our specification of advertising costs (the mirror image of decreasing marginal benefits of advertising is increasing marginal costs of advertising). For this purpose, I include A^2_{it} in advertising costs to allow potentially convex advertising costs.

The period profit of firm i of type L_i is given thus by

$$\pi_{it} = (P_{2it} - MC_{it})M_{it}ms_{it} - (1 + \nu^A_{it})A_{it} - \theta_3 A^2_{it} - \theta_4 1(A_{it} > 0) \tag{8}$$

where P_{2it} is the expense ratio (P_{1it} does not enter the profit function directly since the loads are paid to brokers, not to fund companies), marginal costs $MC_{it} = \theta_1 + \theta_2 L_i + \nu^P_{it}$ is expressed in percentage, ms_{it} is the market share of firm i at time t, and $1(\cdot)$ is an indicator function. In the product market, dominant firms set statically optimal prices under the assumption that they compete in the Bertrand-Nash fashion. Exploiting the first-order conditions of this static pricing game, I recover the parameters of marginal costs of production and the variance of the private MC shocks, θ_1, θ_2, and $\sigma^2_{\nu^P}$.[23]

I have three policy functions to estimate in the first step of BBL procedure: advertising, entry, and exit policy functions. These reduced-form policy functions are estimated as a function of state variables. Essentially, it amounts to regressing observed actions (ad spending, number of entrants, binary decision of exit) on state variables. The idea is that since what we observe in the data is an equilibrium play by firms and this play is based on firms' correct beliefs about their competitors' actions given the state, the optimal policy function can be implicitly expressed as a function of state variables only, albeit a very complicated function. This suggests that flexible functional forms would be desirable for policy function

[22] Firms also incur fixed production costs to operate their business. Fixed production costs include salaries of fund managers and analysts, the rental costs of offices and computer equipment necessary for operation. I do not include fixed costs in the period profit function because (a) for dominant firms, there is no exit and as a result fixed costs are not relevant for their problem, and (b) for fringe firms, fixed costs of operation would not be separately identified from sell-off values.

[23] Since I do not model fringe firms' pricing decisions, I cannot recover fringe firms' marginal costs of production. To compute period profit for fringe firms, I assume that fringe firms have the same marginal cost structure as the one recovered for dominant firms.

estimation. At the same time, however, some degree of parameterization is a must in my application since market size grows over time (although it eventually stops growing) and I need to be able to project the optimal policy functions for market sizes unobserved in the data. In estimating the reduced-form advertising policy function, I exclude 2001-2004 since these years correspond to the period of the dot com bubble burst, which led to drastically different advertising choices for mutual fund companies, as is clear in Table 6.

Advertising policy function for dominant firm i will be modeled as a function of the following state variables; i's goodwill stock, other dominant firms' goodwill stocks, i's quality μ_{it}, other dominant firms' qualities, market size, and the number and average quality of fringe firms. I estimate separate advertising policy functions for load firms and no-load firms, as firms' optimal advertising choices could quite differ depending on their distribution channels. Entry policy function would predict the number of entrants for each type in each period, and I use ordered probit for this, as in Bresnahan and Reiss (1991). As is typical in the two-step estimation literature, I use probit for incumbent fringe firms' exit policy function, and again I do estimation separately for each type.

In the second step of the estimation, I recover the remaining parameters, (θ_3, θ_4, $\sigma^2_{\nu^A}$, $F^L_\phi(\phi)$, $F^{NL}_\phi(\phi)$, $F^L_\kappa(\kappa)$, and $F^{NL}_\kappa(\kappa)$). To estimate these parameters, I use forward simulation proposed by BBL. I simulate the behavior of firms under the optimal strategy recovered in the first step (which I denote σ) and compare the net present value arising from it to the value a firm would obtain if it deviates from the optimal strategy, while its competitors continue to follow the strategy. For the strategy recovered from the first step to be optimal, the value from the former should be higher than the value from the latter. In other words, σ should satisfy

$$V(x_i, \mathbf{x}^D, \overline{\mathbf{x}}^F, z|\sigma_i, \sigma_{-i}, \lambda^L, \lambda^{NL}) \geq V(x_i, \mathbf{x}^D, \overline{\mathbf{x}}^F, z|\sigma'_i, \sigma_{-i}, \lambda^L, \lambda^{NL}) \text{ for } \forall\ i,\ (\mathbf{x}^D, \overline{\mathbf{x}}^F, z), \text{ and } \sigma'_i \tag{9}$$

By finding parameter values that minimize the violation of the inequalities, I obtain a minimum distance estimator. Since the profit function is linear in parameters, I can achieve significant savings in computational costs, as explained by BBL. From the forward simulation, we obtain estimates of (θ_3, θ_4, $\sigma^2_{\nu^A}$, $F^L_\phi(\phi)$, $F^{NL}_\phi(\phi)$). Once I obtain these estimates, I can construct the value function for each incumbent at a given state. Exploiting the fact that the entry cost κ does not enter the value function of incumbents, I recover the distribution of entry sunk costs for each type of potential entrants using:

$$
\begin{array}{c}
\Pr(\lambda^L = N) = \\
F^L_\kappa(\beta E[V(x^e_i, \mathbf{x}^D, \overline{\mathbf{x}}^F, z|\sigma, N, \lambda^{NL})]) - F^L_\kappa(\beta E[V(x^e_i, \mathbf{x}^D, \overline{\mathbf{x}}^F, z|\sigma, N+1, \lambda^{NL})]) \\
\Pr(\lambda^{NL} = N) = \\
F^{NL}_\kappa(\beta E[V(x^e_i, \mathbf{x}^D, \overline{\mathbf{x}}^F, z|\sigma, \lambda^L, N)]) - F^{NL}_\kappa(\beta E[V(x^e_i, \mathbf{x}^D, \overline{\mathbf{x}}^F, z|\sigma, \lambda^L, N+1)])
\end{array}
\tag{10}
$$

$\Pr(\lambda^L = N)$ and $\Pr(\lambda^{NL} = N)$ are quantities observed in the data, and $E[V(\cdot)]$ can be computed once forward simulation is done and $(\theta_3,\ \theta_4,\ \sigma^2_{\nu^A},\ F^L_\phi(\phi),\ F^{NL}_\phi(\phi))$ are estimated. Once we impose a parametric assumption on F^L_κ and F^{NL}_κ (for instance, log normal), we can use (10) to recover parameters characterizing the distribution of entry sunk costs. I treat discount factor β as known and fix it at 0.9 for estimation.

6 Results

Demand estimates are reported in Table 8. The first column in Table 8 reports logit estimates and the second column reports IV estimates, where prices are treated as potentially endogenous. From the estimates, we observe that β_1 is estimated to be positive and significant, while β_2 is estimated to be negative and significant. β_1 and β_2 are parameters on the log of goodwill stock: β_1 is the parameter for no-load firms, and $\beta_1 + \beta_2$ is the parameter for load firms. Therefore, the estimates confirm the prior that consumers who buy products from no-load firms tend to be more advertising sensitive than consumers who buy products from load firms. This difference in consumers' responsiveness to advertising would lead to different advertising behavior for load and no-load firms, as we will see below in policy function estimates. The retention rate of goodwill stock δ is estimated to be 0.6 in the second column. This suggests that advertising loses approximately 40% of its effect after one year, 64% loss after two years, 78% loss after three years, and almost 90% loss after four years. The fact that advertising has only slight impact on goodwill stock after four years provides an ex-post justification for the finite lag distribution used in the goodwill accumulation function (6).

P_1 is the magnitude of loads (commissions), and we see that the size of loads has almost no effect on consumer demand. This might look surprising because one would expect that investors prefer those that charge lower loads among load firms, ceteris paribus. The insignificant coefficient on P_1 could be interpreted as an indication of agency problems between investors and brokers, as discussed in Section 5.2. Although investors do not like higher loads, investors who buy load funds tend to follow brokers' recommendations in making their investment choices. Since loads are what brokers get paid for their advice, brokers might have an incentive to steer investors into higher-load products. P_2 is expense ratios, and the estimates indicate that higher expense ratios reduce consumer utility. Comparing the first and second columns, we see that the impact of expense ratios on consumer utility becomes much larger once we instrument for prices, which is in line with previous findings in the demand literature (e.g., Berry, 1994). The implied own-price elasticity of demand with respect to expense ratios increases from -0.736 to -2.184 once we instrument for prices.

Other demand coefficients are overall plausible. Consumers prefer to buy products from firms that have been around longer, as firm age could be a signal of firm quality and credibility. Consumers also prefer firms that have had good performance in previous years, consistent with the established findings in the mutual fund literature about performance-chasing investors (e.g., Chevalier and Ellison, 1997; Berk and Green, 2004). Somewhat surprisingly, $ln(N)$, included as a regressor following Ackerberg and Rysman (2005), is estimated to be insignificant despite the large number of firms in my application.

Next, I discuss estimates of state transitions. For the transition of goodwill stock, there is no extra estimation to be done, since the retention parameter δ is already recovered from demand estimation. Demand estimates indicate that goodwill stock evolves as follows.

$$G_{it} = \sum_{k=0}^{4} 0.6^k A_{it-k} \tag{11}$$

Once I obtain demand estimates, I can recover quality μ_{it} of each firm from them. Since I model dominant firms' advertising and pricing decisions and μ_{it} is assumed to evolve according to an exogenous process, μ_{it} for dominant firm i should be *net of* goodwill and price effects. Hence, we have $\mu_{it} = \gamma_1 L_i + \gamma_2 X_{it} + \xi_{it}$ for dominant firms. For fringe firms, I do not model their advertising or price decisions. However in the data, they have non-zero prices and might have non-zero advertising as well, although much smaller than dominant firms' ad spending. Therefore, their goodwill, if any, and price effects should be included in μ_{it}. In other words, fringe firms' goodwill and price effects will be assumed to evolve according to an exogenous process along with other firm characteristics, all summarized in the single-index quality measure μ_{it}. Hence, we have $\mu_{it} = (\beta_1 + \beta_2 L_i)\ln(G_{it}) - \alpha P_{it} + \gamma_1 L_i + \gamma_2 X_{it} + \xi_{it}$ for fringe firms. Using the recovered μ's, I then estimate AR(1) transition process for firm quality. I do separate estimation for dominant firms and fringe firms, and also for load firms and no-load firms. I also estimate transition for the average of fringe firms' quality $\overline{\mu}_t^F$ for each type. I need to estimate the transition of this mean quality of the fringe because that's what firms will use in computing their value function under Behavioral Equilibrium (3). The estimated parameters for the transition functions are reported in Table 9. It is clear from the table that there is a very high degree of persistence in firm quality. Finally, estimated transition of market size is reported in Table 9.

Once the demand model is estimated, I can also estimate marginal cost parameters from the optimal pricing game among dominant firms. In the product market, firms set statically optimal prices under the assumption of Bertrand-Nash competition. Since the pricing game is static, this estimation can be done separately from the estimation of the dynamic game. According to the definition of Behavioral Equilibrium, firms do not track the exact distribution of fringe firms' state variables. Instead, they simply condition their strategies on summary statistics of fringe firms' state variables: the number of fringe firms, N_t^F, and their average quality, $\overline{\mu}_t^F$ for the load and no-load types. Hence, in solving for the optimal price P_{2it} (expense ratios), each dominant firm would behave as if it is competing against N_t^{LF} load fringe firms whose quality levels are $\overline{\mu}_t^{LF}$ and N_t^{NLF} no-load fringe firms whose quality levels are $\overline{\mu}_t^{NLF}$, in addition to 29 other dominant firms.[24] Exploiting the first-order conditions of this static pricing game and using the price elasticity estimated from the demand model, I recover the production marginal cost parameters θ_1 and θ_2. The marginal cost parameters are estimated as follows (standard errors inside the parentheses).

$$\begin{aligned} MC_{it} &= 0.0006 + 0.0044 L_i \\ &\quad\;(0.0002)\quad(0.0003) \end{aligned} \tag{12}$$

The estimated parameters suggest that for no-load firms marginal costs are about 0.06% of managed assets, while for load firms marginal costs are about 0.5% of managed assets. These figures might seem low, but recall that unlike typical manufacturing industries the marginal costs of selling one additional share of mutual funds does not involve any additional materials. In this industry, marginal costs include the costs of managing extra accounts, such as providing investors with quarterly summary of their accounts, and possibly incentive pay for fund managers, as managers might get paid more for managing larger funds. For load firms, however, there is an additional cost. Because load firms sell their products through brokers, they need to hire wholesalers to pitch their products to brokers. Load firms give

[24] In demand estimation, I proceeded as if consumers make choices based on the entire distribution of dominant and fringe firms' characteristics. It might seem strange to argue that firms use summary statistics of the fringe in their value function calculation due to "behavioral" reasons while consumers use the whole distribution. However, one can justify this modeling choice by interpreting BE as econometricians' approximation. It is difficult for econometricians to solve the complicated dynamic game with many firms. Hence, we econometricians solve the model as if firms condition their strategies on summary statistics of the fringe. On the other hand, demand estimation is not any more difficult when we have more firms, hence there is no need for approximation.

incentive pay to wholesalers if they sell more funds, so the marginal costs of distributing funds for load firms include these additional pays, and thus are higher.

The last component in the first step estimation is estimation of reduced-form policy functions. First-stage policy function estimates—advertising, number of entrants, and exit—are reported in Table 10. For the advertising policy function of dominant firms, I use a tobit model. As the dependent variable, I use normalized advertising spending ($\frac{\text{AD}}{\text{Market Size}}$) instead of original ad expenditures for a better fit. To be flexible in estimation, I include polynomial terms of the state variables and their interactions. For the entry policy function, I use an ordered probit model. The dependent variable is the number of entrants in a given year for each type. Since the number of observations for the entry policy function is very small, I am forced to use a very parsimonious specification. Finally for the exit policy function of fringe firms, I use a probit model.

Obviously, it is hard to interpret the coefficients in the policy functions because they are intended to be a reduced-form representation of the solution to complicated dynamic games. Hence, I do not attempt to provide structural interpretations of the coefficients. However, I make a few points about the policy function estimates. First of all, we see that no-load firms' advertising spending is more responsive to market size than load firms' advertising spending is. This is what we would expect, given that no-load consumers are more responsive to advertising than load consumers are as the demand estimates suggested. Moreover, advertising spending of no-load firms with large goodwill stocks tends to increase with market size at a faster rate than that of no-load firms with small existing goodwill stocks, while for load firms, the responsiveness of ad spending to market size does not significantly vary with firms' goodwill stocks. This pattern is indicative of an escalation in endogenous sunk costs for those no-load firms who already have a prominent position in the market. We see that less fringe firms enter the market if the goodwill stocks of dominant incumbents are higher, possibly because competition for consumers gets tougher for new fringe entrants if incumbents have accumulated large goodwill stocks. Also, more fringe firms enter the market as the increase in market size is larger. Fringe firms with greater quality are less likely to exit. Overall, the fit of the reduced-form policy functions is fine.

It is crucial to get the policy function estimates right in the first stage estimation, because it is going to be a main driver of market share dynamics in the second stage estimation and in counterfactual analyses. This is true in all problems, but especially important in my setting because I study a growing market and therefore would need to extrapolate the policy functions to states that are not observed in the data. Hence, for my analysis to be valid the relationship between advertising and market size recovered from the data should carry over to market sizes that are unobserved in the data. This issue is not unique to my application and arises in any application where some states are not realized in the data or the state vector does not exhibit a stationary Markov transition (e.g., growing network in Ryan and Tucker, 2008).

Because it is important to get the policy functions right, I do the following exercise. I estimate the reduced-form policy functions in the first stage, and then simulate how the industry will evolve over time under the estimated policy functions using the state vector at the beginning of the sample period as a starting point. Then I compare the observed evolution of the industry in the data to the simulated paths to check how the policy functions perform. Figure 1A shows market share dynamics for no-load and load firms in the data. It shows dominant firms only. From the figure, we see why C_3 and C_5 grew over time in the no-load segment. Two firms kept increasing their market shares, two firms continuously lost their market shares, and the rest of the dominant no-load firms increased their market shares a bit on average. In the load segment, the share of the largest firm remained the same in the beginning and at the end of the sample period, although the identities changed. For all other firms, their shares got smaller on average. Figure 1B shows market share dynamics for no-load and load firms according to the simulated paths under estimated policy functions. Again, it shows dominant firms only. We see that the model reasonably gets the overall pattern right in the sense that no-load firms tend to grow over time, while load firms tend to get smaller. This comes from the fact that in the estimated policy functions, no-load firms' advertising is more responsive to market size. However, the model does not do a good job in explaining asymmetries among firms within a given type. It predicts lower dispersion of market shares within each type than what's actually observed in the data. This suggests that I need to incorporate richer firm heterogeneity, and this is an extension I am currently working on.

The results presented so far summarize first step estimates: demand parameters, transitions for state variables, production marginal cost parameters, and reduced-form policy functions. With these estimates in hand, I perform forward simulation to recover the remaining structural parameters of the model: advertising costs, distribution of advertising cost shocks, and the distribution of sunk costs and sell-off values. This is the second step of BBL estimation.

To perform forward simulation, we exploit the optimality condition of the observed actions being Behavioral Equilibrium, as in (9). To obtain alternative behavioral policies σ_i', I perturb σ_i by adding a random shock ε drawn from a normal distribution with a chosen variance, $\sigma_i' = \sigma_i + \varepsilon$. In conducting forward simulation, I need to convert the reduced-form advertising policy function obtained in the first stage (advertising decision as a function of observed states s_t only) into the true policy function (advertising decision as a function of observed states s_t *and* private advertising cost shock ν_{it}^A). To do so, I exploit the fact that the period profit function has increasing differences in $(A_{it}, -\nu_{it}^A)$. Using this monotonicity and the normal distribution assumption for the advertising cost shock ν_{it}^A, I can back out the optimal policy function from the distribution of ads conditional on s_t, which is estimated as a tobit in the first step.

[FORWARD SIMULATION CURRENTLY IN PROGRESS]

I plan to perform various counterfactual analyses once I estimate structural parameters from the second step. In the meantime, I consider a "counterfactual" scenario which can be examined without structural parameter estimates. In particular, I ask how market structure would evolve over time if no firm is allowed to advertise in this industry. This exercise would shed light on the role of advertising in driving the different market share dynamics between the two segments. In this scenario, firms still differ depending on their quality μ_{it}. In other words, even if no firm advertises, some firms could get ahead if they get successively positive shocks to their quality. Figure 2 shows the resulting market share dynamics under this scenario. It again shows dominant firms only. I use the state vector of year 1989 as a starting point for the industry. Hence, some firms start with high levels of goodwill stocks. However, since there is no additional advertising the initial goodwill stocks will completely disappear after 4 years. That is why Figure 2 shows big drops in market shares of dominant firms in the first few years. Figure 2 clearly shows that without advertising the market shares of load firms and no-load firms who were dominant at the beginning of the sample period would become smaller and smaller as market size grows (and there will be more and more fringe firms as market grows). Without advertising, there are no endogenous sunk costs, so market structure would get fragmented with an increase in market size, and firms who were initially dominant would not be distinguishable from fringe firms eventually. The market share dynamics in the absence of advertising suggests that advertising is an important strategic tool that keeps a concentrated market structure in a growing market.

7 Conclusion

TO BE COMPLETED.

References

[1] Ackerberg, Daniel (2001), "Empirically Distinguishing Informative and Prestige Effects of Advertising," *RAND Journal of Economics*, 32 (2), 316-333.

[2] Ackerberg, Daniel, and Marc Rysman (2005), "Unobserved Product Differentiation in Discrete Choice Models: Estimating Price Elasticities and Welfare Effects," *RAND Journal of Economics*, 36 (4), 771-788.

[3] Aguirregabiria, Victor, and Pedro Mira (2007), "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, 75 (1), 1-53.

[4] Bajari, Patrick, Lanier Benkard, and Jonathan Levin (2007), "Estimating Dynamic Models of Imperfect Competition," *Econometrica*, 75 (5), 1331-1370.

[5] Benkard, Lanier (2004), "A Dynamic Analysis of the Market for Wide-Bodied Commercial Aircraft," *Review of Economic Studies*, 71 (3), 581-611.

[6] Beresteanu, Arie, and Paul Ellickson (2007), "The Dynamics of Retail Oligopoly," working paper.

[7] Bergstresser, Daniel, John Chalmers, and Peter Tufano (2006), "Assessing the Costs and Benefits of Brokers in the Mutual Fund Industry," working paper.

[8] Berk, Jonathan, and Richard Green (2004), "Mutual Fund Flows and Performance in Rational Markets," 112 (6), 1269-1295.

[9] Berry, Steven (1994), "Estimating Discrete-Choice Models of Product Differentiation," *RAND Journal of Economics*, 25 (2), 242-262.

[10] Berry, Steven, James Levinsohn, and Ariel Pakes (1995), "Automobile Prices in Market Equilibrium," *Econometrica*, 63 (4), 841-890.

[11] Berry, Steven, and Joel Waldfogel (2006), "Product Quality and Market Size," working paper.

[12] Bresnahan, Timothy (1992), "Sutton's Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Concentration," *RAND Journal of Economics*, 23 (1), 137-152.

[13] Bresnahan, Timothy, and Shane Greenstein (1999), "Technological Competition and the Structure of the Computer Industry," *Journal of Industrial Economics*, 47 (1), 1-40.

[14] Bresnahan, Timothy, and Peter Reiss (1991), "Entry and Competition in Concentrated Markets," *Journal of Political Economy*, 99 (5), 977-1009.

[15] Chevalier, Judith, and Glenn Ellison (1997), "Risk Taking by Mutual Funds as a Response to Incentives," *Journal of Political Economy*, 105 (6), 1167-1200.

[16] Ciccotello, Conrad, Jason Greene, and Lori Walsh (2005), "Supermarket Distribution of Open-End Mutual Funds," working paper.

[17] Collard-Wexler, Allan (2006), "Plant Turnover and Demand Fluctuations in the Ready-Mix Concrete Industry," working paper.

[18] Cronqvist, Henrik (2006), "Advertising and Portfolio Choice," *Journal of Financial Economics*, forthcoming.

[19] Doganoglu, Toker, and Daniel Klapper (2006), "Goodwill and Dynamic Advertising Strategies," *Quantitative Marketing and Economics*, 4 (1), 5-29.

[20] Doraszelski, Ulrich, and Sarit Markovich (2007), "Advertising Dynamics and Competitive Advantage," *RAND Journal of Economics*, 38 (3), 557-592.

[21] Doraszelski, Ulrich, and Mark Satterthwaite (2007), "Computable Markov-Perfect Industry Dynamics: Existence, Purification, and Multiplicity," working paper.

[22] Dubé, Jean-Pierre, Günter Hitsch, and Puneet Manchanda (2005), "An Empirical Model of Advertising Dynamics," *Quantitative Marketing and Economics*, 3 (2), 107-144.

[23] Ellickson, Paul (2004), "Supermarkets as a Natural Oligopoly," working paper.

[24] Ellickson, Paul (2007), "Does Sutton Apply to Supermarkets?" *RAND Journal of Economics*, 38 (1), 43-59.

[25] Elton, Edwin, Martin Gruber, and Christopher Blake (2006), "The Adequacy of Investment Choices Offered by 401(k) Plans," *Journal of Public Economics*, 90 (6-7), 1299-1314.

[26] Ericson, Richard, and Ariel Pakes (1995), "Markov-Perfect Industry Dynamics: A Framework for Empirical Work," *Review of Economic Studies*, 62 (1), 53-82.

[27] Gallaher, Steven, Ron Kaniel, and Laura Starks (2006), "Madison Avenue Meets Wall Street: Mutual Fund Families, Competition and Advertising," working paper.

[28] Gowrisankaran, Gautam, and Robert Town (1997), "Dynamic Equilibrium in the Hospital Industry," *Journal of Economics & Management Strategy*, 6 (1), 45-74.

[29] Gremillion, Lee (2005), *Mutual Fund Industry Handbook : A Comprehensive Guide for Investment Professionals* (Wiley).

[30] Hendel, Igal (1999), "Estimating Multiple-Discrete Choice Models: An Application to Computerization Returns," *Review of Economic Studies*, 66 (2), 423-446.

[31] Hendel, Igal, and Aviv Nevo (2006), "Measuring the Implications of Sales and Consumer Inventory Behavior," *Econometrica*, 74 (6), 1637–1673.

[32] Jain, Prem, and Joanna Shuang Wu (2000), "Truth in Mutual Fund Advertising: Evidence on Future Performance and Fund Flows," *Journal of Finance*, 55, 937-958.

[33] Kim, Kyoo il, and Minjung Park (2008), "Industry Dynamics in Growing Markets," mimeo.

[34] Krusell, Per, and Anthony Smith, Jr. (1998), "Income and Wealth Heterogeneity in the Macroeconomy," *Journal of Political Economy*, 106 (5), 867-896.

[35] Macieira, Joao (2006), "Extending the Frontier: A Structural Model of Investment and Technological Competition in the Supercomputer Industry," working paper.

[36] Maskin, Eric, and Jean Tirole (1988), "A Theory of Dynamic Oligopoly, I: Overview and Quantity Competition with Large Fixed Costs," *Econometrica*, 56 (3), 549-569.

[37] Pakes, Ariel, and Paul McGuire (1994), "Computing Markov-Perfect Nash Equilibria: Numerical Implications of a Dynamic Differentiated Product Model," *RAND Journal of Economics*, 25 (4), 555-589.

[38] Park, Minjung (2008), "An Empirical Two-Sided Matching Model of Acquisitions: Understanding Merger Incentives and Outcomes in the Mutual Fund Industry," working paper.

[39] Pesendorfer, Martin, and Philipp Schmidt-Dengler (2004), "Identification and Estimation of Dynamic Games," working paper.

[40] Pozen, Robert (2002), *The Mutual Fund Business* (MIT Press).

[41] Reid, Brian, and John Rea (2003), "Mutual Fund Distribution Channels and Distribution Costs," *Investment Company Institute Perspective*, 9 (3), 1-19.

[42] Reuter, Jonathan, and Eric Zitzewitz (2006), "Do Ads Influence Editors? Advertising and Bias in the Financial Media," *Quarterly Journal of Economics*, 121 (1), 197-227.

[43] Roberts, Mark, and Larry Samuelson (1988), "An Empirical Analysis of Dynamic, Nonprice Competition in an Oligopolistic Industry," *RAND Journal of Economics*, 19 (2), 200-220.

[44] Ryan, Stephen (2006), "The Costs of Environmental Regulation in a Concentrated Industry," working paper.

[45] Ryan, Stephen, and Catherine Tucker (2007), "Heterogeneity and the Dynamics of Technology Adoption," working paper.

[46] Shaked, Avner, and John Sutton (1983), "Natural Oligopolies," *Econometrica*, 51 (5), 1469-1483.

[47] Shaked, Avner, and John Sutton (1987), "Product Differentiation and Industrial Structure," *Journal of Industrial Economics*, 36 (2), 131-146.

[48] Sutton, John (1991), *Sunk Costs and Market Structure: Price Competition, Advertising, and the Evolution of Concentration* (MIT Press).

[49] Sutton, John (2006), "Market Structure: Theory and Evidence," working paper.

[50] Weintraub, Gabriel, Lanier Benkard, and Benjamin Van Roy (2008a), "Markov Perfect Industry Dynamics with Many Firms," *Econometrica*, 76 (6), 1375-1411.

[51] Weintraub, Gabriel, Lanier Benkard, and Benjamin Van Roy (2008b), "Extensions to Oblivious Equilibrium," working paper

Table 1
Assets Under Management ($ billion)

Year	Load Segment	No-Load Segment
1985	437.47	311.46
1986	655.53	426.63
1987	665.29	454.08
1988	679.47	475.79
1989	768.02	557.23
1990	830.59	535.68
1991	973.03	655.53
1992	1103.41	777.64
1993	1352.60	950.73
1994	1370.95	951.44
1995	1663.48	1228.67
1996	1923.69	1505.40
1997	2294.14	1928.47
1998	2741.90	2355.15
1999	3275.01	2851.16
2000	3289.60	2734.43
2001	3298.93	2607.13
2002	2976.79	2408.24
2003	3292.62	2833.40
2004	3367.65	3120.83

In 1998 dollars

Table 2
Market Structure - Dominant Firms

Year	Load Segment		No-Load Segment	
	C_3	C_5	C_3	C_5
1985	33.00	44.76	37.83	48.41
1986	33.18	45.94	40.75	51.05
1987	29.97	43.51	40.88	52.25
1988	28.22	42.50	41.26	53.22
1989	28.11	41.47	44.00	56.07
1990	26.08	37.53	51.12	58.38
1991	24.98	35.72	52.92	60.52
1992	24.42	34.66	51.16	59.21
1993	23.17	33.34	51.76	60.13
1994	23.43	33.30	53.51	61.81
1995	23.08	32.80	53.50	61.80
1996	24.15	33.51	52.80	61.01
1997	24.32	33.54	50.99	58.71
1998	23.48	31.88	51.73	59.74
1999	23.19	31.34	54.53	62.35
2000	20.68	29.10	54.97	63.66
2001	19.31	27.25	53.28	63.75
2002	19.37	27.86	52.63	61.88
2003	21.32	28.94	52.19	60.43
2004	25.15	31.78	52.81	60.39

Table 3
Market Structure - Dominant Firms and Fringe Firms ($ billions)

Year	Load Segment		No-Load Segment	
	Average size of the biggest firms who, combined, have 50% of market share	Average size of the smallest firms who, combined, have 50% of market share	Average size of the biggest firms who, combined, have 50% of market share	Average size of the smallest firms who, combined, have 50% of market share
1985	36.626 (N=6)	1.876 (N=116)	27.478 (N=6)	0.759 (N=193)
1986	56.425 (N=6)	2.330 (N=136)	43.556 (N=5)	0.924 (N=226)
1987	51.623 (N=7)	1.853 (N=164)	47.452 (N=5)	0.914 (N=237)
1988	53.409 (N=7)	1.746 (N=175)	50.638 (N=5)	0.856 (N=260)
1989	58.000 (N=7)	1.967 (N=184)	72.443 (N=4)	1.001 (N=267)
1990	53.261 (N=8)	2.032 (N=199)	91.284 (N=3)	0.976 (N=268)
1991	55.586 (N=9)	2.229 (N=212)	115.641 (N=3)	1.182 (N=261)
1992	57.851 (N=10)	2.385 (N=220)	132.612 (N=3)	1.351 (N=281)
1993	70.719 (N=10)	2.806 (N=230)	164.039 (N=3)	1.575 (N=291)
1994	69.838 (N=10)	2.874 (N=234)	169.698 (N=3)	1.404 (N=315)
1995	78.358 (N=11)	3.325 (N=241)	219.127 (N=3)	1.680 (N=340)
1996	90.532 (N=11)	4.069 (N=228)	264.964 (N=3)	1.874 (N=379)
1997	116.226 (N=10)	5.098 (N=222)	327.774 (N=3)	2.299 (N=411)
1998	127.385 (N=11)	5.704 (N=235)	406.135 (N=3)	2.571 (N=442)
1999	150.600 (N=11)	7.067 (N=229)	518.245 (N=3)	2.843 (N=456)
2000	140.028 (N=12)	7.120 (N=226)	501.007 (N=3)	2.773 (N=444)
2001	139.275 (N=12)	7.364 (N=221)	463.064 (N=3)	2.793 (N=436)
2002	126.650 (N=12)	7.284 (N=200)	422.456 (N=3)	2.647 (N=431)
2003	133.169 (N=13)	8.261 (N=189)	492.946 (N=3)	3.295 (N=411)
2004	144.378 (N=12)	9.290 (N=176)	549.326 (N=3)	3.691 (N=399)

Inside the parenthesis is the number of firms that belong to each category

Table 4
Number of Advertisers vs. Non-Advertisers

	Load Segment		No-Load Segment	
	Adverisers	Non-Advertisers	Advertisers	Non-Advertisers
1985	20	102	23	176
1986	23	119	36	195
1987	34	137	32	210
1988	25	157	30	235
1989	25	166	25	246
1990	28	179	24	247
1991	35	186	36	228
1992	32	198	39	245
1993	36	204	46	248
1994	40	204	50	268
1995	37	215	44	299
1996	50	189	52	330
1997	43	189	56	358
1998	42	204	57	388
1999	47	193	52	407
2000	45	193	37	410
2001	36	197	25	414
2002	67	145	67	367
2003	60	142	59	355
2004	55	133	59	343

Table 5
Ad-Sales Ratio

	Load Segment	No-Load Segment
Top 5	0.521	1.504
Others	0.175	0.316

	Load Segment	No-Load Segment
Top 15	0.365	1.106
Others	0.169	0.297

Table 6
Ad Expenditures ($ thousands)

	Load Segment		No-Load Segment	
	Av. Spending by Big 5	Av. Spending by (Big 5)^c	Av. Spending by Big 5	Av. Spending by (Big 5)^c
1985	366.11	63.86	1252.92	27.51
1986	787.55	72.91	3516.40	87.92
1987	878.62	126.04	4398.76	97.48
1988	1210.09	90.42	5941.99	61.53
1989	2043.10	116.93	11073.17	58.49
1990	2832.83	76.14	15015.46	30.27
1991	3500.50	77.13	12776.95	51.85
1992	3773.72	167.50	15717.99	95.82
1993	4322.38	107.64	7063.19	145.25
1994	5455.81	166.13	8839.70	188.40
1995	3557.49	110.52	6335.29	184.01
1996	2630.73	150.31	9175.71	188.05
1997	4239.10	306.54	9674.95	183.57
1998	5327.54	215.49	13262.68	224.72
1999	2226.62	321.32	12859.21	148.95
2000	3503.59	470.27	15822.06	164.23
2001	511.94	329.84	12703.01	87.93
2002	585.06	290.37	4058.71	88.35
2003	4543.84	152.21	10528.22	67.62
2004	4710.76	221.80	14619.30	92.86

Table 7
Are Big Firms Also Big Ad Spenders?

	Load Segment	No-Load Segment
1985	3	4
1986	2	3
1987	1	3
1988	1	4
1989	3	4
1990	2	5
1991	3	5
1992	3	5
1993	3	4
1994	2	4
1995	3	4
1996	2	4
1997	2	3
1998	2	4
1999	1	5
2000	2	4
2001	0	3
2002	1	3
2003	1	4
2004	2	5

The reported figures are the number of big five firms who are also one of the top ten ad spenders within each segment.

Table 8
Demand Estimates

	Logit	Logit with IV
β_1	0.251 (22.974)	0.247 (18.530)
β_2	-0.161 (-11.332)	-0.153 (-8.483)
δ	0.570 (7.040)	0.602 (5.696)
P_1	-0.008 (-0.587)	0.113 (0.874)
P_2	-50.716 (-29.692)	-150.486 (-8.490)
Load Firm	0.378 (6.235)	0.343 (0.839)
ln(N)	-0.221 (-1.415)	0.247 (1.152)
Age	0.081 (5.920)	0.050 (2.932)
Age^2	-0.001 (-4.977)	-0.0008 (-2.331)
1(Age<2)	-0.625 (-7.553)	-0.613 (-6.217)
1(Age<5)	-0.308 (-4.500)	-0.400 (-4.836)
$Perf_{t-1}$	2.351 (13.445)	0.969 (2.953)
$Perf_{t-1}^2$	-1.231 (-6.576)	0.050 (0.121)
$Perf_{t-2}$	1.985 (10.361)	0.056 (0.167)
$Perf_{t-2}^2$	-0.983 (-4.923)	0.571 (1.497)
No. obs	8177	8177
Adjusted R^2	0.674	0.537
Implied own-price elasticity for P_2	-0.736	-2.184

Inside the parentheses are t-statistics.
In IV logit, the instruments for prices are competitors' average number of fund offerings, interacted with load dummy, competitors' average age, interacted with load dummy, and the square of those terms.

Table 9
State Transition Estimates

$$\mu_{it+1} = -0.953 + 0.778\mu_{it} \text{ for dominant load firms}$$
$$(0.105) \quad (0.021) \quad R^2=0.813$$

$$\mu_{it+1} = -0.831 + 0.854\mu_{it} \text{ for dominant no-load firms}$$
$$(0.137) \quad (0.020) \quad R^2=0.923$$

$$\mu_{it+1} = -0.121 + 0.983\mu_{it} \text{ for fringe load firms}$$
$$(0.049) \quad (0.005) \quad R^2=0.94$$

$$\mu_{it+1} = -0.35 + 0.961\mu_{it} \text{ for fringe no-load firms}$$
$$(0.051) \quad (0.004) \quad R^2=0.913$$

$$\overline{\mu}_{t+1}^{LF} = -2.726 + 0.730\overline{\mu}_{t}^{LF} \text{ for average quality of load fringe}$$
$$(1.602) \quad (0.156) \quad R^2=0.627$$

$$\overline{\mu}_{t+1}^{NLF} = -4.743 + 0.578\overline{\mu}_{t}^{NLF} \text{ for average quality of no-load fringe}$$
$$(2.393) \quad (0.212) \quad R^2=0.364$$

$$\Delta M_{t+1} = 128.10 + 0.808\Delta M_t$$
$$(70.16) \quad (0.097) \quad R^2=0.64$$

Inside the parentheses are standard errors.

Table 10

Policy Function Estimates

Advertising Policy Function Dependent variable: AD_{it}/Market $Size_t$		
	Load firms	No-load firms
$G_{it} / \sum_i G_{it}$	14.121 (18.769)	-51.186 (18.124)
$(G_{it} / \sum_i G_{it})^2$	-45.371 (42.082)	-61.677 (24.041)
$(G_{it} / \sum_i G_{it})^3$	-5.530 (181.706)	99.460 (56.634)
$1000 \times \mu_{it} / \sum_i \mu_{it}$	10.363 (4.893)	6.282 (13.480)
$(1000 \times \mu_{it} / \sum_i \mu_{it})^2$	-4.937 (2.497)	0.827 (6.545)
$(1000 \times \mu_{it} / \sum_i \mu_{it})^3$	1.115 (0.704)	-0.369 (1.603)
$1000 \times (\sum_i \mu_{it} - \mu_{it}) / ((N_t - 1)\sum_i \mu_{it})$	8.654 (21.605)	-21.723 (42.335)
$[1000 \times (\sum_i \mu_{it} - \mu_{it}) / ((N_t - 1)\sum_i \mu_{it})]^2$	-4.836 (11.921)	11.894 (23.619)
$[1000 \times (\sum_i \mu_{it} - \mu_{it}) / ((N_t - 1)\sum_i \mu_{it})]^3$	0.818 (2.174)	-2.453 (4.355)
$[G_{it} / \sum_i G_{it}] \times [1000 \times \mu_{it} / \sum_i \mu_{it}]$	5.376 (9.062)	-3.657 (6.142)
$[G_{it} / \sum_i G_{it}] \times [1000 \times (\sum_i \mu_{it} - \mu_{it}) / ((N_t - 1)\sum_i \mu_{it})]$	-1.940 (9.212)	23.312 (8.291)
$M_t / 10000$	1.804 (1.021)	2.047 (3.105)
$[M_t / 10000] \times [G_{it} / \sum_i G_{it}]$	-0.935 (5.519)	19.944 (5.620)
$[M_t / 10000] \times [1000 \times \mu_{it} / \sum_i \mu_{it}]$	-2.971 (1.633)	-3.757 (3.801)
constant	-9.627 (14.206)	10.164 (29.194)
No. obs	225	126
Pseudo R^2	0.763	0.764

Inside the parentheses are standard errors.

$G_{it} / \sum_i G_{it}$ is firm i's share of total goodwill stock. The summation is over dominant firms.

$\mu_{it} / \sum_i \mu_{it}$ is firm i's share of total firm "quality." The summation is over both dominant and fringe firms.

$(\sum_i \mu_{it} - \mu_{it}) / ((N_t - 1)\sum_i \mu_{it})$ is the average quality share of firm i's competitors. N_t is the tota number of firms (both dominant and fringe)

Table 10 Continued

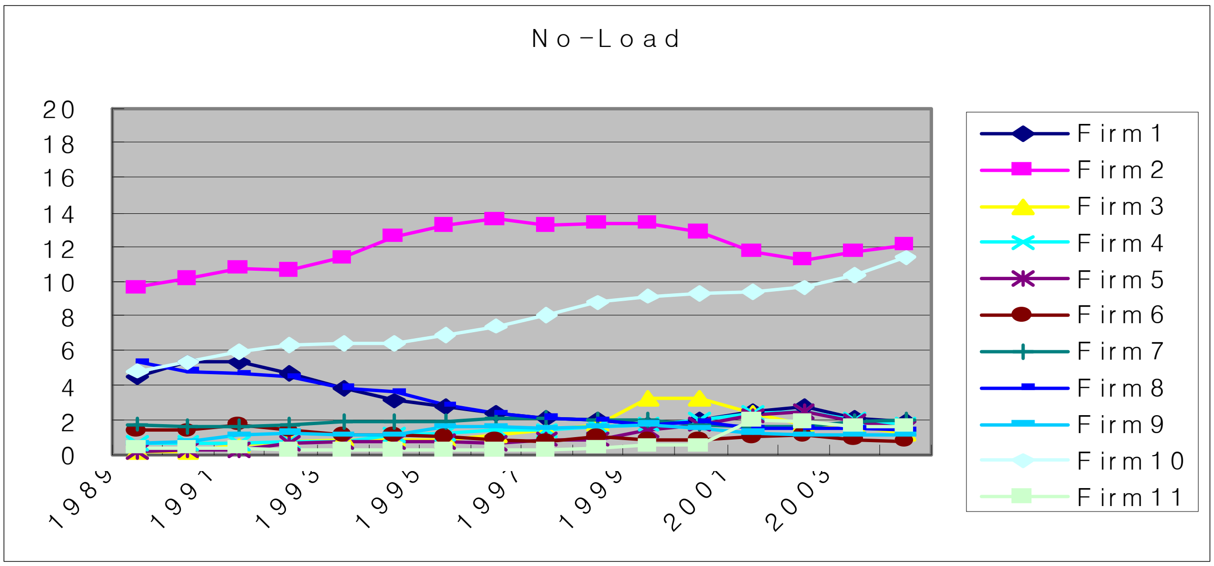
Policy Function Estimates

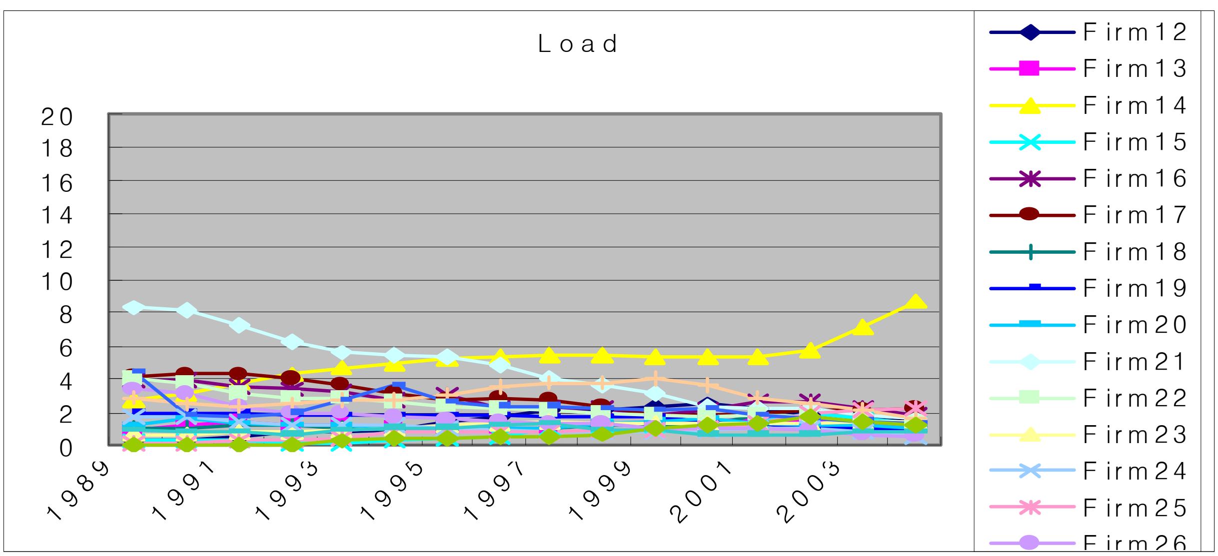
Entry Policy Function Dependent variable: Number of entrants for a given type (load or no-load)	
$\sum_i G_{it}$ / 100000	-1.304 (0.697)
Load type entrants dummy x $\sum_i G_{it}$ / 100000	-1.392 (0.730)
ΔM_t / 10000	1.596 (8.461)
Load type entrants dummy x ΔM_t / 10000	-5.219 (11.953)
No. obs	28
Pseudo R^2	0.192

Exit Policy Function Dependent variable: = 1 if firm i exits in period t, = 0 otherwise		
	Load firms	No-load firms
μ_{it}	-0.098 (0.013)	-0.113 (0.011)
$\sum_i G_{it}$ / 100000	-6.007 (0.491)	-7.476 (0.398)
M_t / 10000	4.960 (0.280)	5.932 (0.231)
ΔM_t / 10000	-12.978 (1.999)	-17.991 (1.559)
$\sum_i \mu_{it}$ (summation over dominant firms only)	0.045 (0.013)	0.076 (0.013)
$\sum_i \mu_{it}$ (summation over fringe firms only)	0.0006 (0.0007)	-0.0004 (0.0006)
N_t^F	0.004 (0.009)	-0.009 (0.008)
constant	2.098 (3.096)	7.851 (2.971)
No. obs	3092	5472
Pseudo R^2	0.323	0.453

Inside the parentheses are standard errors.

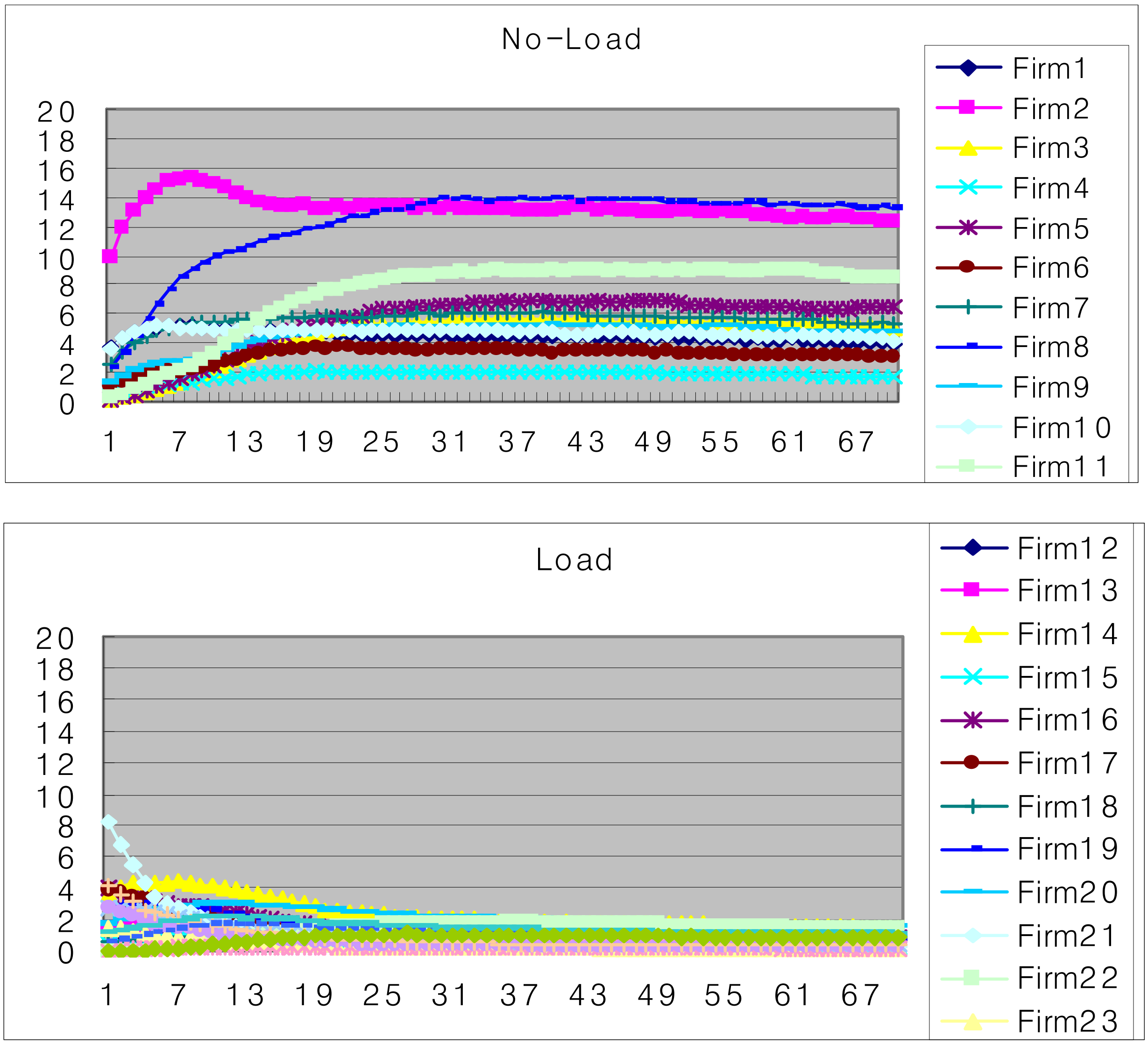
Figure 1A
Market Share Dynamics: Data





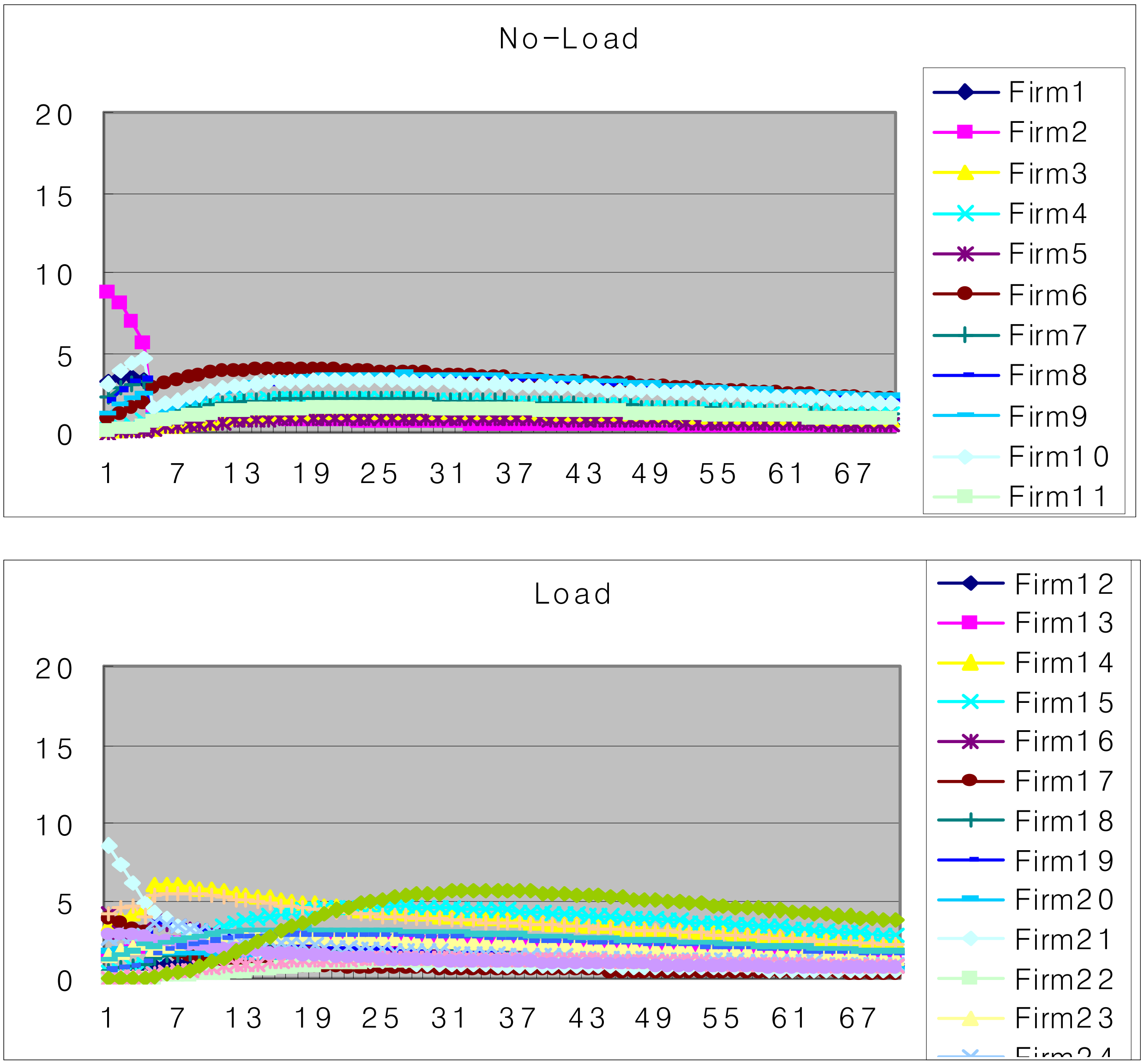
y-axis: market share

Figure 1B
Market Share Dynamics: Model Prediction



y-axis: market share
x-axis: year. 1 correponds to year 1989

Figure 2
Market Share Dynamics: If no firm advertises



y-axis: market share
x-axis: year. 1 correponds to year 1989