Do Search Costs Explain Persistent Investment in Active Mutual Funds?*

Aljoscha Janssen[†] Jurre Thiel[‡]

April 8, 2025

Abstract

Active funds, though losing market share since the 1990s, make up nearly half of all mutual funds but charge more without better performance. We analyze fund data and a search model, highlighting the impact of search costs and active fund preferences. From 1993 to 2018, reduced search costs expanded the market and heightened competition, while a preference shift from active to passive funds increased the latter's market share. However, investors who choose active funds, facing higher search costs, and continue to show a strong preference for them, allow these funds to keep charging higher fees.

JEL: D12, D22, G15, L13

Keywords: Mutual Funds, Search Costs, BLP

*We thank seminar participants at the Copenhagen Business School, the 2021 International Industrial Organization Conference, the Consumer Search Digital Seminar Series, the 2021 Asian Meeting of the Econometric Society, the 2021 Australasian Meeting of the Econometric Society, and the 48th Annual Conference of the European Association for Research in Industrial Economics.

[†]Singapore Management University, ajanssen@smu.edu.sg

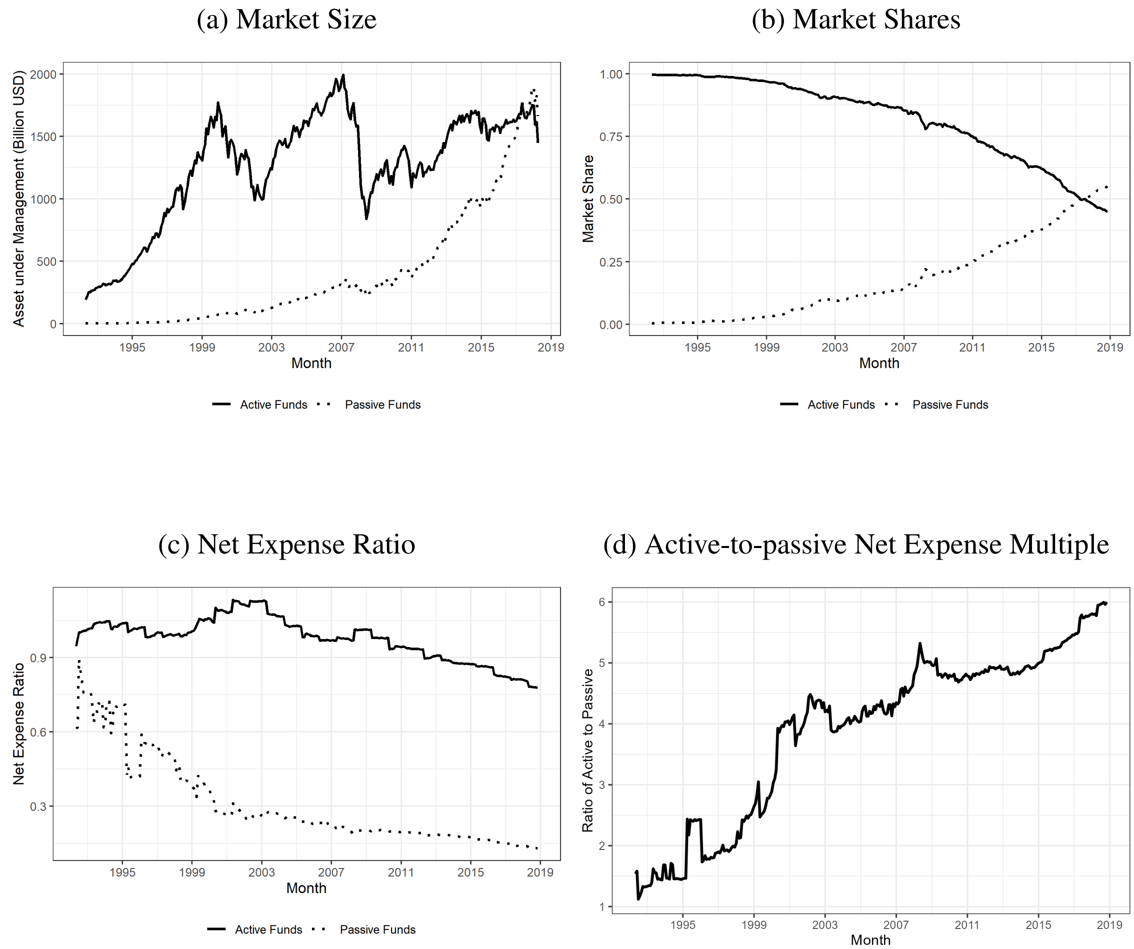
[‡]CPB Netherlands Bureau of Economic Analysis & Vrije Universiteit Amsterdam, j.h.thiel@cpb.nl

1 Introduction

In 2019, 46.4% of American households owned mutual funds (Investment Company Institute, 2020). Within the mutual fund industry, investors must choose between actively or passively managed funds. Since the 1990s, passive index funds have gained popularity. In 1990, less than 3% of managed assets in US equity mutual funds were passively managed, but as of today, the share of passively managed assets has exceeded 50%. This shift in investment strategies is unsurprising, given that net of fees, active funds typically underperform their passive counterparts (Fama and French, 2010, French, 2008, Gruber, 1996, Jensen, 1968). However, a significant portion of investors continue to invest in underperforming, actively managed funds. Moreover, the shift in consumer preferences from active to passive funds has not affected the expenses or performance of active funds. Figure 1 illustrates that, in comparison to passive funds, active funds increased their net expense ratios from the 1990s to the early 2000s, despite losing market share. Since 2003, the relative expenses between active and passive funds have remained stable, indicating that active fund prices seem to be less affected by increased competitive pressure in the 1990s.

The continued strong presence of active funds has been questioned, but there is no single benchmark to define how large an "appropriate" market share for active funds should be. Investors who favor active management might place a high value on perceived skill, trust in fund managers, or additional services, which can justify investing in funds with higher fees. Nevertheless, it is important to understand the forces that keep active funds substantial in size even when passive alternatives are cheaper and often perform better net of fees. This paper focuses on two forces that can jointly explain these outcomes: investors' preferences for active funds and the search costs that they face in collecting and comparing information on multiple funds. If search costs are significant, many investors may remain with higher-fee active funds rather than expending effort to seek cheaper alternatives. At the same time, some investors' willingness to pay for active management may be sufficiently strong, so that they continue choosing active funds even with relatively high prices. Yet, as the market evolves over time and more investors enter, it remains unclear how these preferences and search frictions together explain the shifting shares of active and passive funds.

Figure 1: The Mutual Fund Market



Notes: The figures show (a) monthly market sizes, (b) monthly market shares, (c) average net expense ratios, and (d) the active-to-passive net expense multiple of U.S. equity funds between 1993 and 2018. Mutual fund investments include retail U.S. equity funds from the Morningstar database (Morningstar, 2019). Net expense ratios and average returns are weighted by asset size and reported in percent.

This paper addresses two main research questions. The first question examines whether the persistent high investment in actively managed funds can be attributed to investor preferences, or if it is instead explained by search costs. It is important to distinguish between preferences and search costs in this context because if search costs are the primary driver, there could be implications for policies that aim to improve disclosures and transparency. The second research question seeks to explain the apparent contradiction of increasing competition in the mutual fund market and yet, stable or rising prices of active funds. We use a rich dataset covering all mutual funds in the US equity market between 1993 and 2018 to estimate a structural model that recognizes investors' incomplete information about each fund and their costs of searching. We cannot observe fund-by-fund investments at the individual account level, but we infer how overall inflows and outflows respond to prices and characteristics. In particular, the growing share of passive funds may be partly explained by new, more price-sensitive investors entering the market, while longstanding investors with a preference for active management or higher search frictions stay.

We construct and estimate a discrete choice model in which investors select between various mutual funds and an outside alternative. Building on insights from search theory (Weitzman, 1979; Moraga-González et al., 2017), we show that demand for any particular fund depends on the fund's price, the investor's utility from active or passive management, and a reservation utility that captures how long a search continues. Mutual fund companies choose prices by anticipating these reservation utilities. Our estimation reveals that many investors have a strong preference for active funds, but also face non-trivial search costs. Moreover, we find that search costs have declined by more than 60% over the last three decades, expanding the overall mutual fund market. The influx of new investors who are more price-sensitive has increased competitive pressure, yet active funds continue to rely on a segment of investors with higher search costs or stronger preference for active management.

We then study several counterfactuals to illustrate how search costs and preferences jointly shape market size, fund shares, and prices. If we artificially increase search costs, overall market size decreases substantially, and passive funds lose the most price-sensitive newcomers. If we increase the preference for active funds, overall market size increases, and active funds gain market share, but competition between funds also intensifies. Neither search cost reductions nor preference changes alone fully explains the industry's evolution over time; instead, both are important. The combined reduction in search costs and decreased preference for active funds can explain why passive funds now hold a majority share yet active funds remain in a strong position with relatively high prices.

In the first part of our paper, we shed light on key developments in the mutual fund industry over the last 30 years. We analyze how the market shares and expenses of active and passive mutual funds are related. Our findings reveal several stylized facts: decreasing market shares of active funds, persistent entry of passive funds, high price dispersion between active and passive funds, and a negative correlation between expenses and market share that strengthens for passive but weakens for active. These patterns do not align with simple price competition, especially for active funds, which retain higher prices. In the second part, we develop and estimate our structural search model and discuss how it identifies investor preferences and search costs from equilibrium prices and market outcomes. Finally, we use a series of counterfactual scenarios to decompose the role of search costs versus preferences in shaping market outcomes.

Our paper contributes to understanding one of the most important markets for retail investors, building on the literature streams in both economics and finance. The existing literature has extensively documented the inferior performance of active funds (Fama and French, 2010, French, 2008, Guercio and Reuter, 2014, Gruber, 1996, Jensen, 1968), and many reasons have been proposed to explain why investors continue choosing them, such as investors being naive (Gruber, 1996), broker incentives (Bergstresser et al., 2008), or the "peace of mind" offered by skilled managers (Gennaioli et al., 2015). Our analysis combines these perspectives with the idea of costly search (Hortaçsu and Syverson, 2004; Sirri and Tufano, 1998) to evaluate how investors may fail to fully explore cheaper alternatives. We further extend prior work on search costs and price dispersion (Wolinsky, 1986, Stahl, 1989, Burdett and Judd, 1983, Reinganum, 1979) by examining how search costs shape active and passive competition. Finally, our approach is related to structural estimations of demand with search frictions (Moraga-González et al., 2023; Hortaçsu and Syverson, 2004). Here, our main technical contribution is that we develop a general model of random search that can be estimated using only market-level data, i.e., fund prices and market shares. Previous contributions only encompassed vertical differentiation between products (Hortaçsu and Syverson,

2004) or required individual-level data (e.g., Honka, 2014). Moraga-González et al., 2023 also allow for both these features, but they assume a directed search protocol in which consumers know the price and some characteristics of the product ex-ante and then inspect products that are likely to be good matches. This contrasts with our assumption of random search in which all products are inspected with equal probability. Some markets are better approximated by directed search, while for others random search is more appropriate. Under directed search, consumers are more likely to inspect and purchase low-price products. Since we can observe that the market share of high-priced active funds remains high, we think a random search model is more appropriate for the mutual funds market. This search protocol is also used in earlier work on the mutual funds market (Hortaçsu and Syverson, 2004).

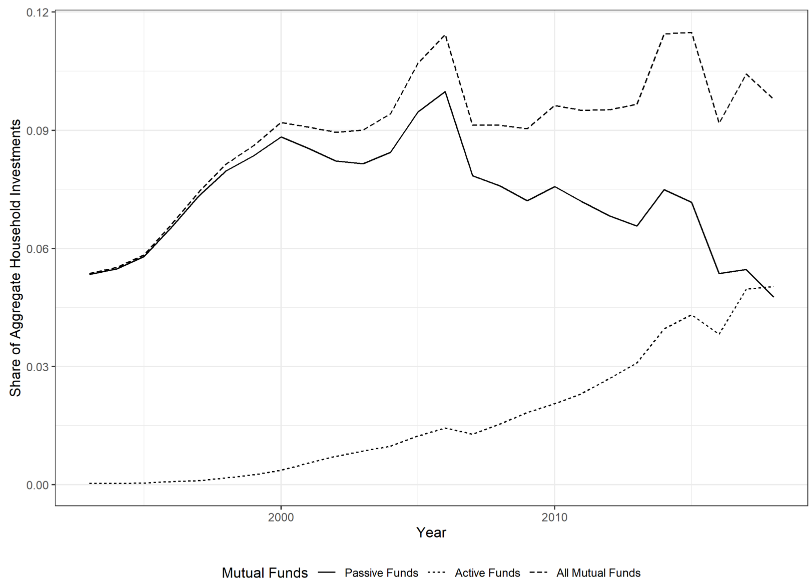
2 Data

We analyze a comprehensive dataset that covers both active and passive mutual funds in US equity markets from 1993 to 2018. To obtain this data, we use Morningstar's fund database (Morningstar, 2019), which includes all US equity funds traded in US dollars, including exchange-traded index funds. Our dataset consists of a monthly panel of fund-level data, including various fund characteristics observed at both monthly and yearly intervals. To focus on the behavior of private investors, we only consider share classes of funds targeted toward retail investors.[1]

A key variable in our analysis is the *net expense ratio*, which we treat as the "price" of a fund. The net expense ratio is the annual percentage of fund assets paid for the fund's operating and management expenses, net of any fee waivers or reimbursements offered by the fund. In practical terms, it incorporates various costs such as accounting, administrative, advisory, auditing, and distribution (12b-1) fees, but excludes brokerage costs or one-time sales charges. We specifically use the *Annual Report Net Expense Ratio* from Morningstar, reflecting the actual fees charged over

[1] We exclude mutual fund classes that typically require minimum investments of more than $50,000 and those that are not directly available to retail investors, such as mutual funds available only through retirement accounts. Thus, we limit our analysis to mutual fund classes A, B, C, T, and No Load.

Figure 2: Share of Mutual Funds Among all Investments



Notes: The Figure presents the yearly share of mutual fund investments among all household investments between 1993 and 2018. Mutual fund investments consider all investments in US equity funds traded in the US, based on Morningstar's mutual fund database (Morningstar, 2019). We obtain the share of mutual fund investments among all assets by using data of financial account data for all households in the United States, provided by the board of governors of the Federal Reserve System (The Board of Governors of the Federal Reserve System, 2019). We aggregate data on the yearly level and show shares of mutual fund investments of all mutual, as well as only active or passive funds.

a given fiscal year. This measure is widely used by both industry practitioners and the academic literature because it directly affects an investor's net returns on a recurring basis.

We connect the data on equity funds and their investments to financial account data for all households in the United States, provided by the board of governors of the Federal Reserve System (The Board of Governors of the Federal Reserve System, 2019). This enables us to compare households' equity investments to other available investment options. Figure 2 shows the yearly share of mutual fund investments considering yearly aggregate financial investments of US households.

Additionally, Table 1 presents summary statistics for both active and passive funds, divided into three time intervals. We observe a substantial increase in the raw number of both active and passive funds from 1993 to 2018. While the number of active funds increased by 63%, the number of passive funds increased by more than 600%. The second panel of Table 1 shows the fund size weighted average returns in percentage. We find that active funds offered marginally higher returns in the 1990s and 2000s, but this relationship reversed in the last ten years. The standard deviation of monthly returns across each fund category indicates that returns are more dispersed with active funds. The best-performing funds across active funds offer superior returns compared to the high-return passive funds. In panel C, we see the development of monthly average market shares for active and passive funds. As previously presented in Figure 1, active funds lose market share while passive funds gain market share, leading to a less concentrated market over time. Notably, we observe a decreasing Herfindahl-Hirschman Index (HHI) over time for passive funds, indicating a trend towards a less concentrated market. Additionally, on the firm level, we see that the market becomes less concentrated for active funds as the leading financial institution loses market share over time. However, for passive funds, the concentration on the top is the opposite, as the leading firm has increased its market share from 1% in the 1990s to 35% between 2010 and 2018. Finally, panel D of Table 1 shows the fund size weighted yearly net expense ratios of each mutual fund type. We find that active funds increased their expenses from the 1990s to the early 2000s but decreased their expenses afterward. In contrast, passive funds have consistently decreased their net expense ratios over time. Aggregate decreases in prices are much larger for passive compared to active funds.

3 The Mutual Fund Industry

In this section, we conduct an empirical analysis of the market and present our findings on the key developments of market shares and net expense ratios in the industry, as illustrated in Figure 1. Our analysis reveals that while passive funds continue to gain market shares, active funds consistently lose them. Somewhat surprisingly, we find that net expense ratios do not exhibit a clear correlation

Table 1: Summary Statistics

Table	Active Funds 1993-2000	Active Funds 2001-2009	Active Funds 2010-2018	Passive Funds 1993-2000	Passive Funds 2001-2009	Passive Funds 2010-2018
A: Number of Funds						
Avg. Number of Funds	1622	3703	2639	58	268	412
	(728)	(186)	(287)	(45)	(46)	(60)
Avg. Number of Mngmt Firms	156	267	342	5	22	46
	(38)	(28)	(13)	(3)	(4)	(11)
B: Returns						
Avg. Monthly Returns	1.29	0.18	0.93	1.27	0.14	1
	(3.84)	(4.75)	(3.73)	(4.11)	(5.24)	(3.68)
Standard Dev. Monthly Returns	2.1434	1.7097	1.2526	1.133	1.6585	1.0112
	(0.0312)	(0.03)	(0.0267)	(0.1336)	(0.1478)	(0.0731)
Monthly Returns top 10%	8.32	8.39	7.81	7.85	8.08	7.1
	(3.4)	(2.55)	(2.93)	(3.86)	(2.46)	(1.1)
C: Competition						
Market share	0.98	0.88	0.65	0.02	0.12	0.35
	(0.01)	(0.04)	(0.1)	(0.01)	(0.04)	(0.1)
HHI in Percent	1.84	1.22	1.31	14.04	10.56	5.78
	(1.84)	(1.22)	(1.31)	(14.04)	(10.56)	(5.78)
Market share Top Mngmt Firms	0.23	0.15	0.09	0.01	0.06	0.09
	(0.01)	(0.03)	(0.01)	(0)	(0.03)	(0.02)
D: Costs						
Net Expense Ratio	1.02	1.04	0.89	0.55	0.24	0.18
	(0.02)	(0.06)	(0.05)	(0.14)	(0.03)	(0.02)

Notes: The table presents basic summary statistics of monthly data of US equity based mutual funds between 1993 and 2018. The first three columns consider active funds while the second three columns show result for passive funds. In each fund category we divide our data in three time brackets: 1993 to 2000, 2001 to 2009, and 2010 to 2018. A Management Firm refers to the owner of a fund which may emit multiple funds. HHI is the abbreviation for the Herfindahl-Hirschman Index, which is the sum over all monthly market shares. Thereby, the index measures the degree of competition in the market. Standard deviations are calculated over monthly variation and are reported in parentheses.

with market share trends for active funds. Specifically, we observed two distinct periods for active funds: between the 1990s and 2000s, net expense ratios slightly increased, while from the end of the 2000s to 2018, they decreased. In contrast, passive funds consistently decreased their net expense ratios between 1993 and 2018. Additionally, we present a hypothesis that rationalizes these developments and demonstrate that these stylized facts are also observable on an individual fund level.

We analyze the relation between the market share and expenses of fund i in month t in the following two regression models:

$$Share_{it} = \beta_1 ExpenseRatio_{it} + \beta_2 ExpenseRatio_{it} \cdot PassiveFund_i + \gamma \mathbf{X_{it}} + \rho_i + \tau_t + \varepsilon_{it} \quad (1)$$

$$Share_{it} = \beta_3 ExpenseRatio_{it} + \beta_4 ExpenseRatio_{it} \cdot Post2003_t + \gamma \mathbf{X_{it}} + \rho_i + \tau_t + \varepsilon_{it}, \quad (2)$$

where $Share_{it}$ represents the market share of a fund, while $ExpenseRatio_{it}$ denotes the net expense ratio of the fund, which changes annually rather than monthly. Additionally, we use a dummy variable $PassiveFund_i$ that takes the value of one for passively managed funds and zero for actively managed funds. Control variables X_{it} include past performance measures such as the fund's return over the previous month $(t-1)$ and year. We introduce fund-specific fixed effects ρ_i and year-month fixed effects τ_t sequentially. This approach allows us to evaluate the correlation between net expense ratios and market shares for active and passive funds while controlling for year-month fixed effects at the individual fund level. Specifically, the first equation tests the correlation between net expense ratios and market shares for active and passive funds by introducing variation on an individual fund level, while controlling for year-month fixed effects.

Our second regression model focuses solely on active funds and aims to investigate whether there was a shift in the relationship between net expense ratios and market share in 2004, as suggested by Figure 1. To test this hypothesis, we introduce a dummy variable $Post2003_t$ that takes the value of one for the months after 2003. Specifically, we expect a positive coefficient β_4 if the net expense ratio has a greater effect on the market share of a fund, as higher expense ratios may have a greater impact on reducing market share prior to 2003.

Our results are presented in Table 2, where sub-specifications (1) to (4) correspond to the first regression model, and sub-specifications (5) and (6) to the second model. Across all four sub-specifications of the first model, we find a negative correlation between annual net expense ratios and market shares of funds, independent of control variables or fixed effects. Interestingly, this negative correlation is even stronger for passive funds, with sub-specifications (3) and (4) demonstrating that the impact of an increased net expense ratio is at least ten times higher for passive funds than for active funds. Turning to the second regression model, which focuses on active funds only, sub-specifications (5) and (6) show that the net expense ratio is significantly negatively correlated with market share before 2003. However, after 2003, this correlation decreases, and an F-test of sub-specification (6) indicates that we cannot reject the hypothesis of non-existent correlation between net expense ratio and market share for active funds after 2003.[2] Therefore, we conclude that, after controlling for fund fixed effects and time shocks, a negative correlation between expenses and market shares only exists before 2004. This observation aligns with the descriptive findings of Figure 1, which do not show a clear relationship between net expenses and market shares of active funds.

In the previous regression models, we assumed a linear relationship between the net expense ratio and the market share of funds over time. However, it is possible that the correlation changes for either active or passive funds. To investigate this, we ran a simple regression of the market share on the expense ratio for each year and for active and passive funds separately ($Share_{it} = \alpha_0 + \beta_1 ExpenseRatio_i + \varepsilon_{it}$). Figure 3 displays the coefficient estimates $\hat{\beta}_1$ for each fund type and year. Interestingly, for passive funds, we observe a clear linear trend of decreasing coefficients, indicating a more negative correlation over time. On the other hand, for active funds, we observe

[2] We formally test if $\beta_3 + \beta_4 = 0$, with the test statistic of 0.1174 leading to a failure to reject the null hypothesis of no correlation.

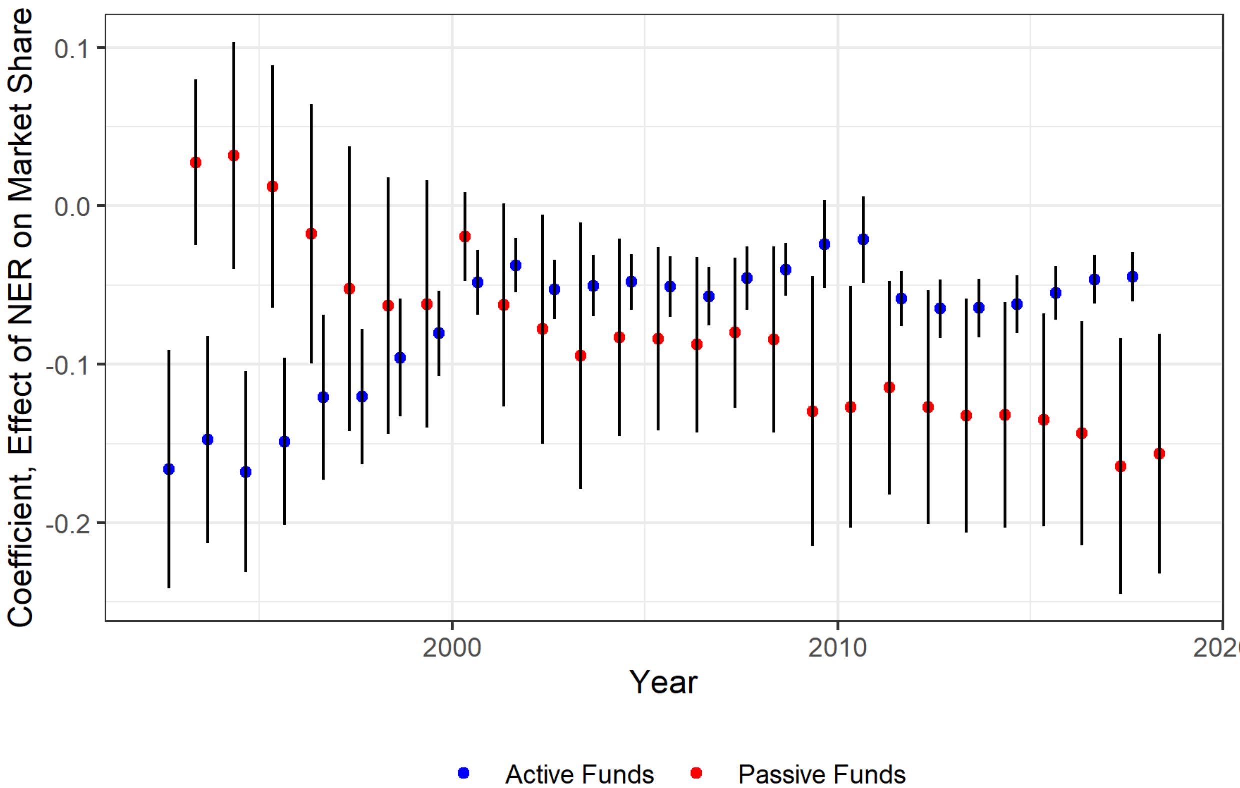
Table 2: Regression, Correlation between Expense Ratio and Market Share

	Market Share					
	All Mutual Funds				Active Funds	
	(1)	(2)	(3)	(4)	(5)	(6)
Expense Ratio	-0.059^{***}	-0.012^{***}	-0.006	-0.012^{**}	-0.107^{***}	-0.034^{***}
	(0.012)	(0.004)	(0.004)	(0.005)	(0.020)	(0.009)
Passive Fund	0.024	0.053				
	(0.045)	(0.054)	(0.000)	(0.000)		
Expense Ratio · Passive Fund	-0.056^{*}	-0.103^{***}	-0.085^{**}	-0.095^{**}		
	(0.034)	(0.036)	(0.036)	(0.044)		
Post2003					-0.151^{***}	
					(0.034)	
Expense Ratio · Post2003					0.064^{***}	0.030^{*}
					(0.016)	(0.017)
Constant	0.134^{***}					
	(0.023)					
Month FE	No	Yes	Yes	Yes	No	Yes
Fund FE	No	No	Yes	Yes	No	Yes
Fund specific Controls	No	Yes	No	No	No	No
Past Return Controls	No	Yes	No	Yes	No	Yes
N	874,206	800,494	874,206	800,649	799,559	731,516

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: The Table shows results for two different regression models presented in 1 and 2. One observation corresponds to a mutual fund in a month. In sub-specifications (1) to (4) we include all mutual funds, in (5) and (6) we reduce the sample to only active funds. The Expense Ratio shows is the yearly net expense ratio of a fund. Passive Fund is a dummy that takes the value one if a fund is passively managed. Post2003 is a dummy that takes the value one for months after 2003. Year-month FE and Fund FE show if fixed effects are included. Fund-specific controls include the tenure of a fund, the category (defined by Morningstar, i.e. large growth, mid-cap, S&P tracking, aggressive allocation, etc.), the Sharpe ratio over the last year, the turnover ratio, the strategic beta, the management company, and the equity style. Past returns control indicates if we control for returns of the past quarter and year. Note that the sample size in the model decreases as we do not observe the controls for all funds. Standard errors are clustered on the fund level, adjusted for heteroskedasticity and serial correlation, and noted in parentheses.

Figure 3: The Relation between Prices and Market Demand



Notes: The Figure presents coefficients of a regression of the market share on the net expense ratio, individually for each fund type (active and passive) and year ($Share_{it} = \alpha_0 + \beta_1 ExpenseRatio_i + \varepsilon_{it}$). The coefficient shows the correlation between an increased expense ratio on a market share of a fund. The error bars correspond to the 95% interval.

the opposite: the relationship between the net expense ratio and market share becomes less negative over time.

Our empirical analysis provides a descriptive account of the observed correlations between net expense ratios and market shares, and does not establish causality. However, we offer a possible explanation for our findings based on a simple framework. We argue that technological advancements and innovations over time have reduced the search costs for investors in finding and comparing investment options. This reduction in search costs has made investors more sensitive to prices, leading to increased competitive pressure in the market. In the early years of our sample, active funds dominated the market. Lower search costs may increase market size, and new price-sensitive consumers entered. Additionally, some consumers change from active to passive funds. Overall we see competitive pressure in the market with decreasing prices for passive funds.

However, active funds, which continued to attract investors with strong preferences, responded by holding their prices constant to exploit remaining consumers to maximize profits. As technology continued to reduce search costs, the market became even more competitive, resulting in lower expenses for investors overall. Nevertheless, the effect on active and passive funds was different. Low-search-cost investors continued to flock to passive funds, while high-search-cost investors remained in active funds, leading to a weaker correlation between expenses and market share for active funds, as they were less price-sensitive. In contrast, the correlation between expenses and the number of low-search-cost investors became stronger for passive funds (see Figure 3).

The simple framework assumes that search costs decrease over time and a sufficient part of investors have stayed with active funds due to search costs or preferences. In the following analysis, we build a model of search that incorporates individual-specific search costs and preferences for active funds.

4 A Model of Search

In the following section, we introduce a model of consumer behavior in the context of mutual fund investment.

Model Setup. We consider T markets, where each market corresponds to a quarter. The size of each market is denoted by M_t, which represents the total financial wealth of households in that quarter. We assume that each dollar of investable wealth corresponds to a unique "consumer" in our model. The set of available mutual funds for consumers in market t is denoted by $\mathscr{J}_t$, and the number of available funds is denoted by $|\mathscr{J}_t|$.

Consumer preferences. Consumer preferences follow a random-utility specification:

$$u_{ijt} = X'_{jt}\beta_{it} - \alpha_{it} p_{jt} + \xi_{jt} + \varepsilon_{ijt} \quad \text{for } i \in 1, \ldots, M_t, j \in \mathscr{J}_t, t \in 1, \ldots, T.$$

Here, u_{ijt} denotes the utility consumer i in market t derives from product j. X_{jt} contains K observable product characteristics for fund j in market t. We assume that X_{jt} contains only variables that differ between products. The price of mutual fund j in market t is denoted by p_{jt}. The price of a mutual fund typically consists of multiple components. To get a single "price", we use a fund's net expense ratio, which aggregates these components in a standard way.[3] The econometric error term is $\xi_{jt} + \varepsilon_{ijt}$. The first component of this term, ξ_{jt} contains the unobserved (to the econometrician) product quality of product j in market t. ε_{ijt} contains idiosyncratic errors, which we assume to be i.i.d. across consumers, products and markets. α_{it} and β_{it} are coefficients we will estimate. We take the following random coefficient specification for these coefficients:

$$\begin{pmatrix} \alpha_{it} \\ \beta_{it} \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix} + \eta_{it},$$

where η_{it} is a $K+1$-dimensional random variable with mean zero. We can also write

$$u_{ijt} = \delta_{jt} + \mu_{ijt} + \varepsilon_{ijt},$$

where $\delta_{jt} = X'_{jt}\beta - \alpha p_{jt} + \xi_{jt}$ is the utility component common to all consumers, while $\mu_{ijt} = (1; -p_{jt}; X_{jt})'\eta_{it}$ is consumer-specific.

Although we do not observe individual investor flows or track specific portfolios over time, our approach infers the evolution of the mutual fund market through the notion of an "outside good." In our model, the outside good corresponds to all alternative household investments beyond the set of mutual funds under consideration. As total investments in mutual funds expand or contract relative to the outside good, we interpret this as respectively capturing net inflows of new (often

[3] The net expense ratio is calculated as the ratio of a fund's operational costs and the fund's net assets.

more price-sensitive) investors or outflows of existing investors. Thus, even without account-level panel data, the demand system with an outside good allows us to *indirectly* gauge the entry of new investors into the market and the retention of others in active funds.

If a consumer buys the outside option, it holds one dollar of financial wealth in another asset than mutual funds.[4] We denote the outside option with $j = 0$ and make the standard normalizations $\delta_{0t} = \mu_{i0t} = 0$ for all i, t. In other words, the utility of the outside option is

$$u_{i0t} = \varepsilon_{i0t} \quad \text{for } i \in 1, \ldots, M_t, t \in 1, \ldots, T.$$

Consumer search. The environment so far is the discrete choice setting that is standard in much of the empirical industrial organization literature (Berry et al., 1995, BLP from now on). We depart from this setting by relaxing the assumption that consumers have perfect information. Instead, consumers engage in costly search to find their best match.

Throughout the article, we treat "search cost" and "cost of information acquisition" as interchangeable constructs. In practice, "search cost" can capture everything from the time it takes to read a disclosure document to the fees associated with professional advice. An investor's decision to invest in an active fund depends on whether the expected benefit of further searching (or learning about a fund manager's skill) outweighs the associated cost of collecting that information. If search costs decrease—due, for instance, to improved online comparison tools—then more investors become aware of passive funds' advantages, potentially fueling a shift away from expensive active funds.

Formally, we model search as random and sequential, consistent with the literature on mutual funds (Hortaçsu and Syverson, 2004). In our market, products are highly substitutable, so directed search would likely lead to fierce competition (Choi et al., 2018), which seems inconsistent with the high prices observed. Random search means that consumers do not observe any component of utility u_{ijt} before they inspect fund j. Consequently, they cannot direct their search toward funds

[4]Our measure of financial assets contains deposits, credit market instruments, corporate equities, security credit, life insurance reserves, pension fund reserves, investment in bank personal trust and equity in noncorporate business besides mutual fund shares.

that are likelier to offer more utility. Therefore, every fund is inspected with equal probability. Search is sequential, and consumers inspect funds one at a time. Consumers always observe their value of the outside option, u_{i0t}, for free. To learn the utility of fund j, consumer i in market t incurs an additively separable search cost of s_{ijt}.

$$s_{it} = s_{it}^{c} + s_{t}^{m}, \tag{3}$$

where s^c is a *c*onsumer-specific search cost and s^m is *m*arket-specific. We assume that search costs are the same for all products: it is not possible to separately to disentangle vertical differentiation from product-specific sampling probabilities with market-level data (Hortaçsu and Syverson, 2004). In our specification, we include firm fixed effects in X_{jt}. These will reflect unequal sampling probabilities, for example stemming from advertising, as well.

Contrary to most of the previous literature, we do not assume that the first search (for an inside good) is free. As a result, a change in search costs not only affects how much consumers search, but also how many and which consumers search (Moraga-González et al., 2017).

Consumers' optimal search is as follows. First, denote by u_{it} the utility obtained when drawing a random inside good. The cdf of u_{it} is hence

$$F(u_{it}) = \frac{1}{|\mathscr{J}_t|} \sum_{j \in \mathscr{J}_t} F(u_{ijt}),$$

where $F(x)$ denotes the cdf of a random variable x. Then, define consumer i's reservation value $\hat{u}_{it}$ as the solution to

$$s_{it} = \int_{\hat{u}_{it}}^{\infty} (u_{it} - \hat{u}_{it}) dF(u_{it}). \tag{4}$$

The left-hand side is the cost of inspecting fund j. The right-hand side is the expected gain in utility if a consumer's best option so far delivers utility $\hat{u}_{it}$. The reservation utility is thus the utility at which the consumer is exactly indifferent between inspecting product j and consuming its best-inspected option so far.

Consumers' optimal search rule follows from Weitzman (1979). Since all products have the same reservation value, consumers are indifferent between visiting any option and sample randomly. The optimal stopping rule is myopic: consumers stop searching as soon as the highest inspected option so far exceeds the reservation value, and keep searching otherwise.

Market shares. We now derive consumers' purchasing probabilities. Say consumer i inspects fund j. Denote by $\mathscr{S}_{ijt}$ all funds he has inspected so far, including j itself and the outside option. The probability that he stops searching and immediately buys product j is

$$P\left(u_{ijt} \geq \max\left\{\max_{k \in \mathscr{S}_{ijt}} u_{ikt}, \hat{u}_{it}\right\}\right) = P\left(u_{ijt} \geq \hat{u}_{it}\right). \tag{5}$$

That is, the consumer must prefer j over all sampled alternatives (so that he buys j) and u_{ijt} must be larger than the reservation value (so that he stops searching). However, a consumer only visits fund j if all previously inspected funds provide utilities below the reservation value. Hence, the requirement $\hat{u}_{ijt} \geq \hat{u}_{it}$ implies that j is preferred over all hitherto sampled alternatives.

It is also possible that the consumer, after inspecting fund j, continues searching and later "comes back" to purchase j anyway. The probability of this event is

$$P\left(\hat{u}_{it} > u_{ijt} \geq \max_{k \in \mathscr{J}_t} u_{ikt}\right). \tag{6}$$

The probability follows from the stationarity of the optimal search protocol. As long as the best-sampled option is worse than the reservation value, the consumer keeps searching. This means that a consumer only comes back after sampling *all* options in the market, which happens when $\hat{u}_{it} > \max_{k \in \mathscr{J}_t} u_{ikt}$. Moreover, we need that the consumer does not purchase product j immediately upon inspection ($\hat{u}_{it} > u_{ijt}$) as well as that product j is preferred over the other products in the market ($u_{ijt} \geq \max_{k \in \mathscr{J}_t} u_{ikt}$).

A main innovation of this article is that consumer's eventual purchasing probabilities follow a discrete choice structure when the number of products is large, so that the probability of comebacks in equation (6) becomes small. To be precise, we let the number of products $|\mathscr{J}_t|$ go to infinity, while holding the number of consumers per product, $M_t/|\mathscr{J}_t|$, and the marginal utility of products in market t, u_{it}, constant. Then, if the distribution of u_{ikt} is unbounded,

$$P\left(\hat{u}_{it} \geq \max_{k \in \mathscr{J}_t} u_{ikt}\right)$$

goes to 0, and hence the probability of comebacks in equation (6) goes to zero. In a large market, it is hence possible to approximate purchase probabilities with equation (5). Because in our application, there is an average of over 2300 products per market, we view this as a reasonable approximation.

To see how this assumption leads to a discrete choice-type formulation, note that under this assumption the probability that consumer i purchases product j is simply

$$P(i \text{ buys } j) = P(i \text{ buys } j | i \text{ visits } j) P(i \text{ visits } j).$$

Under our large-market assumption, all consumers that search at least once eventually make a purchase. This is a result of the fact that as the number of products becomes large, the probability that all have utilities smaller than the reservation value goes to zero. Hence,

$$P(i \text{ searches}) = \sum_{k \in \mathscr{J}_t} P(i \text{ buys } k) = \sum_{k \in \mathscr{J}_t} P(i \text{ buys } k | i \text{ visits } k) P(i \text{ visits } k)$$

and we can write

$$P(i \text{ buys } j) = P(i \text{ searches}) \frac{P(i \text{ buys } j | i \text{ visits } j) P(i \text{ visits } j)}{\sum_{k \in \mathscr{J}_t} P(i \text{ buys } k | i \text{ visits } k) P(i \text{ visits } k)}.$$

The visitation probabilities depend on consumer beliefs. Like most of the theoretical search literature, but contrary to Hortaçsu and Syverson (2004), we assume passive beliefs. This means that when a firm charges an off-equilibrium price, consumers do not update their beliefs on the distribution of utilities in the market. As a consequence, firms choose their prices taking the distribution of reservation utilities as given. In equilibrium, these reservation utilities must be consistent with the prices firms actually charge. Because we cannot separately identify beliefs from preferences, we must assume that this is the case in the data.

Assumption 1 (Equilibrium beliefs)**.** *Consumer beliefs are in equilibrium in the data.*

Under Assumption 1, the probability that any particular product is inspected is the same for across products.[5] Hence the final purchasing probabilities are proportional to the purchasing probabilities conditional on inspection:

$$P(i \text{ buys } j) = P(i \text{ searches})\frac{P(i \text{ buys } j|i \text{ visits } j)}{\sum_{k\in\mathscr{J}_t} P(i \text{ buys } k|i \text{ visits } k)} = P(u_{i0t} < \hat{u}_{ijt})\frac{P(u_{ijt} > \hat{u}_{it})}{\sum_{k\in\mathscr{J}_t} P(u_{ikt} > \hat{u}_{it})}.$$

The second term has a similar structure as the logit discrete choice model. We exploit this similarity we exploit below to show that, under appropriate assumptions, market shares coincide with those from the mixed logit model.

By integrating over consumer types, the market share of product $j > 0$ is then

$$\sigma_{jt} = \int_i P(u_{i0t} < \hat{u}_{ijt})\frac{P(u_{ijt} > \hat{u}_{it})}{\sum_{k\in\mathscr{J}_t} P(u_{ikt} > \hat{u}_{it})} dF(\eta_i). \tag{7}$$

We stress that the last expression for the market shares is only valid under Assumption 1. The reason is that an unanticipated price for product j does not impact the probability that it is inspected. However, it does change the probability that every other product is inspected, because consumers that inspect j are more likely to continue searching. For example, a marginal increase in p_j marginally increases the probability of every other product to be inspected. Hence, equation (7) cannot be used to compute optimal prices—for this the demand function in equation (8) should be used. However, the formulation for the market share is useful because it provides a link with the literature on the estimation of discrete choice models.

[5] To see why we need Assumption 1, consider the case where one firm provides more utility than consumers expect, for example by charging a below-equilibrium price. Because this deviation is not observed before search, the visitation probability of the deviating firm is unaffected by the deviation. However, all *other* firms are less likely to be inspected because consumers are more likely to stop searching after visiting the deviating firm.

Supply. We now derive firms' profit functions. To do so, we derive the demand for product j for an arbitrary, i.e. possibly different from consumer expectations, price p_j.

We start by deriving the probability a consumer inspects a particular product. Conditional on searching at least once, this equals

$$\frac{1}{|\mathscr{J}_t|} + \frac{|\mathscr{J}_t|-1}{|\mathscr{J}_t|}\left(P(u_{it} \leq \hat{u}_{it}) + P(u_{it} \leq \hat{u}_{it})^2 + P(u_{it} \leq \hat{u}_{it})^3 + \dots\right),$$

i.e. the probability that it is visited first, second, etc. As the number of products becomes large, this converges to $P(u_{it} \leq \hat{u}_{it})/(1-P(u_{it} \leq \hat{u}_{it}))$.

Because consumers have passive beliefs, the probability that a consumer inspects a particular product is independent from the product's price and firms take the number and distribution of visiting consumers as given. Hence, the only part of demand that a firm can influence through its price is the conditional purchasing probability $P(u_{ijt} \geq \hat{u}_{ijt})$.

The demand for product j is then

$$d_j(p_j, \{\hat{u}_{it}\}) = \frac{M_t}{|\mathscr{J}_t|}\int P(u_{i0t} < \hat{u}_{it})\frac{P(u_{it} \leq \hat{u}_{it})}{1-P(u_{it} \leq \hat{u}_{it})}P(u_{ijt} \geq \hat{u}_{it})dF(\eta_i). \tag{8}$$

The first term is the size of the market per-product. The integrand, taken over consumer types, contains the probability that a consumer searches, that it visits firm j and the probability it purchases upon inspection. We include the distribution of reservation utilities, $\{\hat{u}_{it}\}$, as an argument to $d_j(\cdot)$ to stress that the demand for a given product depends on consumer beliefs.

If we denote the marginal cost of product j with c_{jt}, the profit of firm f is simply

$$\pi_{ft}(p_t, \hat{u}_t) = \sum_{j \in \mathscr{F}_{ft}} d_j(p, \hat{u}_t)(p_{jt} - c_{jt}). \tag{9}$$

Equilibrium. Market equilibrium occurs when

1. Consumer search is optimal and beliefs are correct. Hence, market shares follow from equation (7).

2. Every firm f sets its prices $\{p_{jt}\}_{j \in \mathscr{F}_{jt}}$ to maximize its profits π_{ft} given the prices of the firms and consumer beliefs $\{\hat{u}_{it}\}$.

Interpreting Persistence and Preference. As is common in the literature (e.g., Hortaçsu and Syverson, 2004), our model focuses on a static equilibrium to explain persistent investment in active funds. In reality, other forces such as switching costs or time-varying beliefs could also produce inertia—particularly for investors already holding active funds. However, we can only incorporate one major friction in this setting, and we argue that search costs are more relevant than switching costs for several reasons.

First, the mutual fund market grew substantially over our sample period, implying a continuous influx of new, often price-sensitive, investors. This expansion reduces the scope for a purely "locked-in" group to dominate active-fund holdings, since many entrants can choose cheaper passive funds if they prefer. In the presence of significant switching costs, we might still see high fees for the locked-in incumbents, but the large inflow of new participants should, in principle, drive active-fund fees toward lower levels over time—yet we do not observe substantial convergence in fees.

Second, if switching costs were the main driver of persistence, we would expect more pronounced downward pressure on active-fund fees as passive funds attract a growing share of new inflows. Over time, the pool of locked-in active investors would diminish, forcing active funds to lower fees or risk losing market share. However, active funds have maintained relatively high expense ratios compared to passive funds, suggesting that a distinct subset of investors genuinely prefers active management and/or faces substantial search frictions, rather than remaining solely due to lock-in.

Third, a fully dynamic model that explicitly tracks switching costs would require detailed panel data on individual choices (such as Honka, 2014), which we lack. Our static approach instead bundles various inertia-generating factors into heterogeneous preferences and nontrivial search frictions. Even moderate preferences for active funds, combined with high search costs, can effectively deter investors from searching for cheaper alternatives, thereby creating de facto persistence in active-fund holdings. Consequently, persistent investment in active funds emerges not purely from strong tastes for active management, but also from the difficulty of gathering and interpreting information on alternative funds.

In summary, while future research with investor-level data could refine these insights using a dynamic framework, our analysis already demonstrates that preferences for active funds and significant search costs jointly suffice to explain high fees and a substantial active-fund presence. Importantly, we do not treat persistent investment in active funds as a simple matter of "preference leads to preference." Instead, our model shows that even modest preferences can sustain active funds when search frictions are large, providing a mechanism for observed inertia without relying exclusively on explicit switching costs.

5 Estimation

We now give a brief overview of our estimation procedure. We develop a new approach to estimating structural search models, in which preference parameters are estimated using standard methods and separately from search costs. Full technical details can be found in Appendix B.

We propose a two-step approach to understand how consumers choose among many products when they must search to learn about prices and quality. First, we show that under appropriate parametric assumptions on the idiosyncratic error term ε_{ijt}, the conditional purchase shares (i.e., shares among those who actually search) behave as if they were the usual inside shares in a typical discrete-choice model. Hence, we can use established methods (Berry et al., 1995) to estimate consumers' taste parameters. The key twist is that, since we focus only on inside shares, we do not include the outside option in the same way as a standard discrete choice model and must instead fix a reference product to handle this additional normalization. In addition, we must control for selection into the market: certain types of consumers tend to search more often. To address this, we incorporate a parametric assumption on preference heterogeneity η_{it} that depends on the number of consumers that is searching.

In the second step, we separately recover search costs and firms' marginal costs. Our main insight is that the fraction of people who do not search equals the outside good's share under our large-market assumption. By combining that share with the first-step demand estimates, we can back out average search costs. A crucial piece of identification comes from observing how market shares respond to shifts that affect the value of the outside option differently from search costs (for instance, cyclical demand changes that raise overall investment). On the supply side, we connect firms' profit conditions to their observed prices and recover marginal costs; we separate the effects of search costs from preferences for products by exploiting that only the former influences prices when consumers still have to search.

6 Results

We will now present the results of our search model. To organize our findings, we first report estimates of consumer preferences before moving on to estimates of search costs.

Preference parameters. Table 3 presents the results of the BLP estimation, providing insight into customers' preferences in the mutual fund market. In the first column, we observe a negative price elasticity of demand, indicating that higher prices (measured as net expense ratio) decrease the utility for customers on average. Additionally, we find a negative effect for passive funds, indicating a strong preference for individuals to invest in active mutual funds. While the coefficients for price and the indicator for a passive fund are random and vary across individuals, we also include multiple non-random coefficients such as the Sharpe ratio, tenure of the mutual fund, and past yearly and quarterly returns, as well as fund category dummies. We allow for correlation between the two random coefficients, price and the indicator for a passive fund, and find that an increase in price sensitivity is positively correlated with a stronger aversion to passive funds (covariance of 2.665). Investors with a higher preference for active funds tend to be more price-sensitive.

The second column describes the impact of the share of an outside good. If fewer customers invest in mutual funds, and the market share of the outside good increases, customers are getting less price sensitive. Additionally, a higher market share of the outside good further increases the preferences for active funds. This can reflect an underlying preference but is also consistent with the idea that the preference for active funds picks up some inertia in investing behavior: since new investors are not locked-in to any fund, the fact that an increasing market share coincides with an increased preference for passive funds can be explained by the fact that existing investors faced lock-in at their current fund. Hence, we can to some extent disentangle true preferences for active funds (as given by the mean preference estimate) from possible persistence (through the interaction term with the size of the outside option). Overall the observation shows that additional investment in the markets (and a lower market share of outside goods) increase price-sensitivity and make passive investment more attractive for the average consumer. Finally, in column three, we show the variance of the random coefficients for price and the indicator of a passive fund, providing insight into the heterogeneity of customer preferences in the market.

Search Costs. In the second step, we present the results of the search cost estimation. Figure 4 displays the evolution of search costs over time, with a cubic smoothing spline. The estimated search costs for a one thousand dollar investment were 0.3 basis points in the early 1990s, decreasing to around 0.075 basis points in the 2000s and 2010s. This translates to a decrease in search costs of almost 75% over the sample period. We observe a decrease of just under 8% from the beginning of our sample until 2000, consistent with the hypothesis put forth by Hortaçsu and Syverson (2004) that search costs decreased in the 1990s, a time when the overall mutual fund market experienced strong growth. We are reassured by the search cost estimates, which vary relatively smoothly across quarters without showing extreme fluctuations. The estimated search costs are considerable but not unrealistically high.

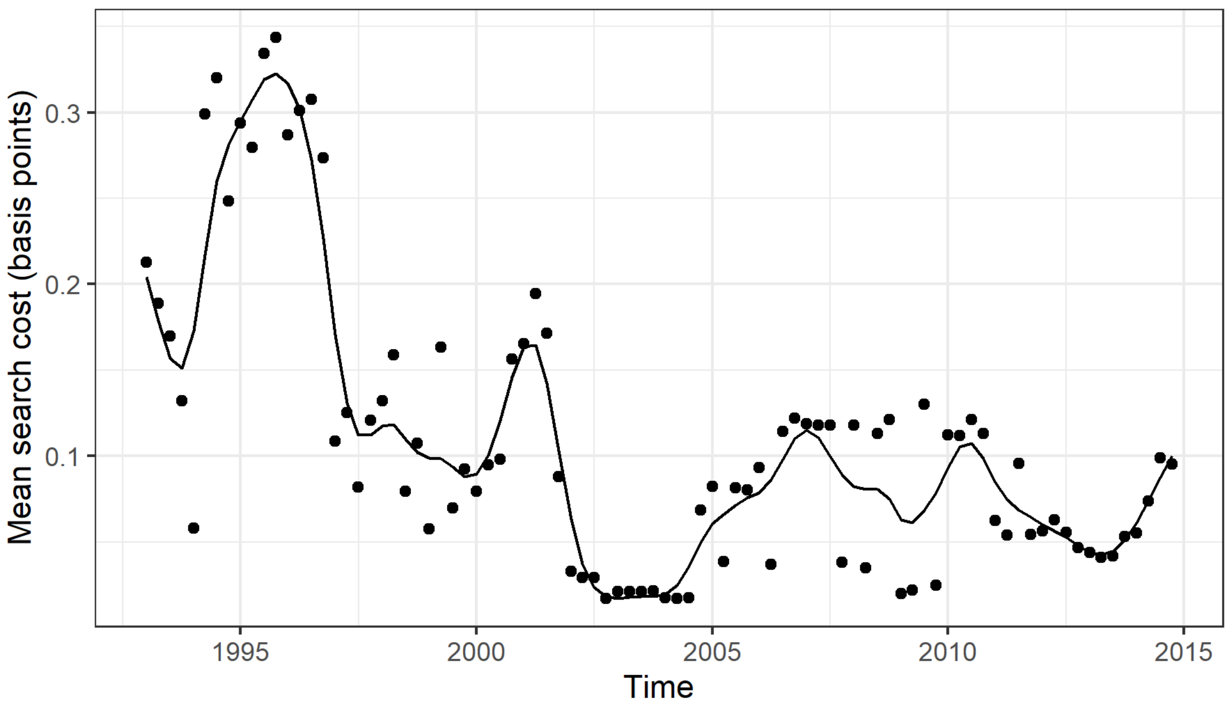
Figure 5 displays the distributions of estimated search costs for consumers who purchase either active or passive funds. We find that there is significant heterogeneity in search costs within both

Table 3: Preference Estimates

	Average Coefficient	Market Share Outside Option	Variance
Price	-5.461	0.268	1.896
	(0.877)	(.021)	(0.758)
Passive Fund	-5.963	-0.108	9.936
	(1.038)	(0.033)	(2.736)
1 Year Sharpe Ratio	-0.046		
	(0.022)		
Tenure of Fund	0.005		
	(0.000)		
Gross Return Past Year	-0.063		
	(0.010)		
Gross Return Past Quarter	0.120		
	(0.026)		
Fund Category Dummies	Yes		
Cov(Price, Index Fund)	2.665		
	(1.601)		

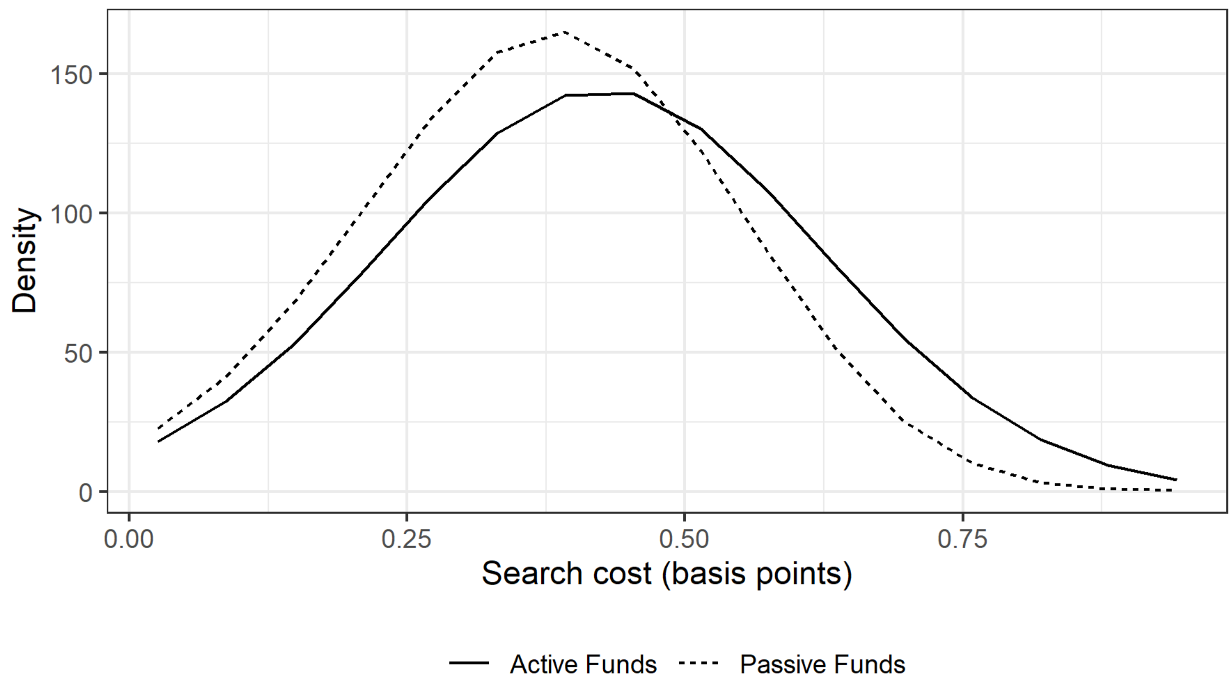
Notes: The table presents estimated preference estimates. The first column displays the average preference for the fund attribute. The second column shows how the preference changes as fewer customers enter the market. This controls for the selection of preferences at the extensive search margin. The third column displays the variance of the preference for the fund attribute across customers. "Cov(Price, Index Fund)" measures the covariance between the random coefficients of these fund attributes. A positive covariance means that The covariance implies a correlation of 0.614. Standard errors are in parentheses.

Figure 4: Search Costs Over Time



Notes: The Figure presents estimated average search costs across quarters. The solid line refers to a cubic smoothing spline.

Figure 5: Search Costs and Density of Mutual Fund Purchases



Notes: The Figure presents the distribution of estimated search costs of those purchasing active or passive funds.

fund types, with search costs ranging from zero to one basis point. The distributions for both types of funds exhibit a longer right tail, indicating that some consumers experience very high search costs, up to 0.8 to 1 basis points. In contrast, there is a clustering of density for both fund types at the lower end of the search cost range.

On average, investors who purchase active funds experience higher search costs than those who purchase passive funds. The distribution of search costs for active funds is shifted to the right. Moreover, among investors with high search costs, in the range of 0.6 to 0.8 basis points, we observe a significantly higher density of active funds compared to passive funds.

The estimation of preferences and search offers five key insights. First, investors prefer active funds compared to passive funds. Second, investors of passive funds are, on average, more price-sensitive. Third, increasing the number of investments, increases price sensitivity and preferences for passive funds. Fourth, investors experience search costs, which are declining between the 1990s and 2010s. Fifth, search costs for investors in active funds are on average larger than those for passive funds.

7 Counterfactuals

While the estimation shows preferences for active funds and large, decreasing, and heterogeneous search costs, it is unknown to which extent each of the factors drives the market shares and prices of active and passive funds. To evaluate the impact of preferences and search costs on market outcomes we perform two counterfactual analyses varying each of the two determinants of choice.

Initially, we select the year 2014 for our counterfactual examination for two primary reasons. This selection aims to curtail the computational demands associated with the counterfactual computation and to isolate the impact of search costs and preferences on market shares. As Figure 4 illustrates, search costs have experienced a decline over time, concurrently with a decrease in the share of active funds. This shift in market shares could be attributed to variations in search costs or might solely reflect changes in market structure. By focusing on a singular year, we aim to replicate the effect of reduced search costs while minimizing the influence of market structure alterations.[6] In examining the market shifts through the lens of increased search costs and a preference for active funds, we aim to reverse-engineer the market dynamics observed in previous years, marked by a surge in passive fund investments and a reduction in prices.

To establish a foundational benchmark, we compute the supply equilibrium based on the demand estimation from the final quarter of our analysis period. The specifics of the counterfactual estimation process are detailed in Appendix C. The baseline of the actual estimation is shown in Table 4.

In the first counterfactual, we increase mean search costs by 20%, while holding the variation

[6] The choice of 2014 is strategic, allowing us to reference historical trends while ensuring the presence of both active and passive funds in the analysis.

across consumers constant. This adjustment implies an escalation in the average search cost for investors, albeit without altering the distributional spread of these costs. The assortment of funds available in the market remains unchanged, permitting investors the liberty to choose, inclusive of the option to abstain from market participation. In the subsequent counterfactual, we enhance the preferences for active funds by 20%. Following the methodology employed in the baseline model, we recalibrate the demand curve, alongside new price points and investor decisions for this modified scenario. The equilibrium outcomes derived from both counterfactual scenarios are documented in Table 4.

Table 4: Counterfactuals: The Effects of Search Costs and Preferences

	Baseline	Increasing Search Costs		Increasing Preference for Active Funds	
Market Size	0.31	0.27	-13.8%	0.34	10.5%
Market Share Active Funds	0.76	0.77	1.2%	0.88	15%
Market Share Passive Funds	0.24	0.23	-3.7%	0.12	-48%
Avg Price Active Funds	0.94	1.00	5.8%	0.91	-2.9%
Avg Price Passive Funds	0.19	0.25	32.2%	0.17	-11.7%

Notes: The table displays counterfactuals that explain the effects of search costs and preferences. The counterfactuals rely on preference and search cost estimates from 2014. The Baseline shows computed quantities from our estimated model without changes in any parameters. In the higher search cost counterfactual, we increase mean search costs by 20%. In the higher preference active funds, we increase the preference for active funds by 20%, increasing the preference for purchasing any fund κ_t *such that the average utility in the market remains constant. For each counterfactual, we show the percent change compared to the baseline. We consider the market size, measured in the ratio of all household assets, as well as the market share and the average price of active and passive funds.*

In the scenario where search costs are elevated, a notable contraction in market size is evident when juxtaposed with the baseline scenario. This phenomenon underscores the pivotal role of search cost reductions in broadening the external margin, thereby incentivizing new investors into the mutual fund sector. A decrease in search costs could catalyze a concurrent expansion in the dimensions of both active and passive mutual funds. By contrast, when search costs rise, passive funds shrink less in absolute size than active funds. As a result, passive funds lose market share while active funds gain. A larger number of investors choosing not to invest indicates a stronger preference for active funds as the marginal investor new in the market prefers passive funds. Therefore, we conclude that higher search costs lead to fewer investors and a slight shift in market shares towards active funds among those who stay.

Looking at fund prices, we found that prices for both passive and active funds went up when search costs were higher. Specifically, prices for passive funds went up by 32.2%, while active fund prices rose by 5.8%. This indicates that search costs have a bigger impact on the prices of passive funds. With higher search costs, some investors decide to leave the market altogether. According to our findings in Table 3, new investors usually prefer passive funds and are more price-sensitive. Thus, it's mostly these cost-conscious passive fund investors who exit the market. With these price-sensitive investors gone, there's less competition on price, leading to bigger price increases for passive funds.

Increasing search costs has a significant impact on the external margin, and some investors leave the market. According to the results in Table 3, new investors tend to prefer passive mutual funds and are more price-sensitive. Thus, we see those customers with preferences for passive funds leave the market. The reduction in competitive pressure is larger for passive funds as their price-sensitive investors leave the market. Thus, we see larger price changes for passive funds.

In our second counterfactual analysis, we explore the effects of enhancing the attractiveness of active mutual funds by 20%. As indicated in Table 4, this heightened appeal results in a 10.5% expansion of the overall market size. This shift precipitates significant alterations in market dynamics, primarily characterized by a substantial migration of investors from passive to active funds. Consequently, the market share of passive funds plummets by 48%, while that of active funds increases by 15%. This trend aligns with the logical expectation that an average shift in investor preference from passive to active funds would reorder market shares accordingly.

Interestingly, the prices for both active and passive funds witness a downturn, with passive funds experiencing a steeper decline. At first glance, this price reduction might appear paradoxical, especially considering that a swing in preference toward active funds might typically predict an uptick in their prices. The explanation for this phenomenon lies in the enlarged market size, which intensifies competition among funds vying for the attention of the price-sensitive investors. These new entrants, in comparison to the incumbent investor base, display a stronger preference for passive funds, thereby exerting downward pressure also on active funds. In essence, the augmentation in preference for active funds culminates in an enriched investor pool, a pronounced pivot towards active funds, and a general compression of fund prices, with passive funds bearing the brunt of this trend.

Combining an increase of search costs and preferences for active funds. We broaden our investigation to assess the concurrent effects of augmenting search costs and preferences for active mutual funds on market dynamics. Using 2014 as our reference year, we strive to gauge the extent to which a blend of changes in preferences and search costs can replicate the market evolution observed up to that point. For context, the market size of mutual funds in 2011 was 17% lower than in 2014, with active funds holding a 13.3% higher market share and passive funds 41% lower. Furthermore, the average prices of active funds were 5.9% higher and those of passive funds 7.8% higher in 2011 compared to 2014. Our model, while not aiming for an exact chronological match, seeks to assess the correlation of these shifts with our simulated outcomes.

We explore scenarios of 5%, 10%, 15%, and 20% increases in both search costs and preferences for active funds, with Figure 6 delineating the results. Specifically, Subfigure 6b examines the impact on market size, revealing that heightened search costs negatively affect market size, whereas elevated preferences for active funds tend to boost investor numbers. When these effects are combined, the deterrent effect of increased search costs predominates, albeit the enhanced preference for active funds does slightly bolster the investor base, which predominantly favors active funds, counteracting the market contraction to some extent.

Subfigure 6b also details how these adjustments influence the market share of active funds. Notably, an uptick in active funds' market share inversely affects that of passive funds, and vice versa.

Solely increasing search costs marginally benefits active funds' market share, while amplifying the preference for active funds significantly boosts their market dominance. Increased search costs tend to drive out more passive investors, yet the overall shrinkage in market size may diminish absolute investments in both fund types. A simultaneous rise in search costs and preferences for active funds markedly enhances active funds' market share, synergizing the individual effects of each factor.

Regarding pricing, Subfigures 6c and 6d contrast the impacts on active and passive funds respectively.[7] Rising search costs elevate prices for both fund types, with passive funds more acutely affected due to the exacerbated reduction in market competition. Conversely, escalating preferences for active funds tend to lower prices across the board, attributed to an enlarged market and the influx of price-sensitive investors. Merging increased preferences for active funds with higher search costs generally sees the latter's influence dominate, leading to overall price increases. This outcome aligns with the observed market contraction under heightened search costs, reinforcing the intertwined relationship between market size, competition, and pricing dynamics.

Next, we juxtapose the model's forecasted alterations with the actual shifts observed from 2011 to 2014. In Figure 7, our objective is to gauge the extent of discrepancy between the model's predictions—stemming from various combinations of search cost adjustments and active fund preference modifications—and the tangible changes recorded between 2011 and 2014. It's acknowledged upfront that a perfect alignment between predictions and actual data is not anticipated; however, the analysis elucidates that neither shifts in active fund preferences nor search cost alterations singularly account for the observed industry dynamics.

Specifically, Figure 7 delineates the percentage point difference in the predicted expansion of market size, active funds' market share, and the average pricing for active and passive funds in each counterfactual scenario relative to the actual evolution noted from 2011 to 2014. As we increase search costs and preferences for active funds, we essentially evaluate the reversed growth

[7] We acknowledge minor non-linearities in price equilibria for increasing search costs which elude precise explanation, yet we remain confident in the general pricing trends for both fund categories.

from 2014 to 2011 and compare the difference in model prediction to data. Subfigure 7a focuses on market size, Subfigure 7b on active funds' market share, and Subfigures 7c and 7d on the pricing of active and passive funds, respectively. The analysis underscores that the isolated elevation of active fund preferences or search cost increments insufficiently encapsulates the sector's progression. A synthesis of enhancements in both dimensions is necessitated for a comprehensive explanation.

Considering the market size, the empirical data imply pronounced search cost escalations from 2014 to 2011 (or reductions from 2011 to 2014). Nevertheless, the solitary increase in search costs falls short of elucidating the surge in active funds' market share or the average price elevation of passive funds. It's only with the amplification of active funds' preferences do we observe an uptick in their market share, accompanied by relatively modest price hikes in passive funds.

Reflecting on the mutual fund industry's development between 2011 and 2014, notable observations include an enlarged market size, a significant downturn in active funds' market share, and a more pronounced price reduction in passive versus active funds. Our counterfactual analysis thus intimates the necessity for search cost reductions to facilitate market size growth, alongside a diminution in active funds' preferences to counterbalance the market size impact on active funds' market share. Ultimately, the combined effect of decreasing active fund preferences and reduced search costs offers a plausible explanation for the more substantial price decrease in passive funds relative to the moderate price adjustments in active funds.

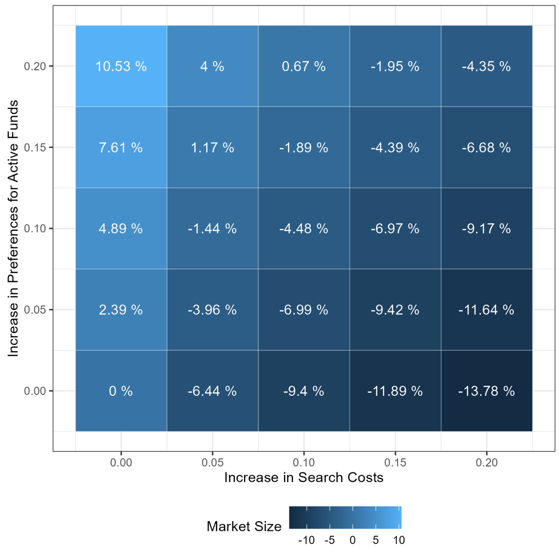
The counterfactual analysis enriches our understanding of the model's estimations by elucidating the effects of altering search costs and preferences for active mutual funds. It demonstrates that elevating search costs significantly diminishes the external margin, catalyzing a withdrawal of investors from the mutual fund market. This withdrawal notably affects the pricing dynamics of passive funds more severely, attributed to the loss of the marginal investor who is more prince-sensitive. Consequently, a reduction in market size disproportionately inflates the prices of passive funds due to previously higher competitive pressures.

Conversely, bolstering preferences for active funds not only amplifies their market shares but also induces a general decrease in fund prices across the board, facilitated by the expansion of market size. This dynamic underscores the intricate balance between investor preferences and market forces, revealing how shifts in favor towards active funds can influence pricing structures throughout the mutual fund landscape.

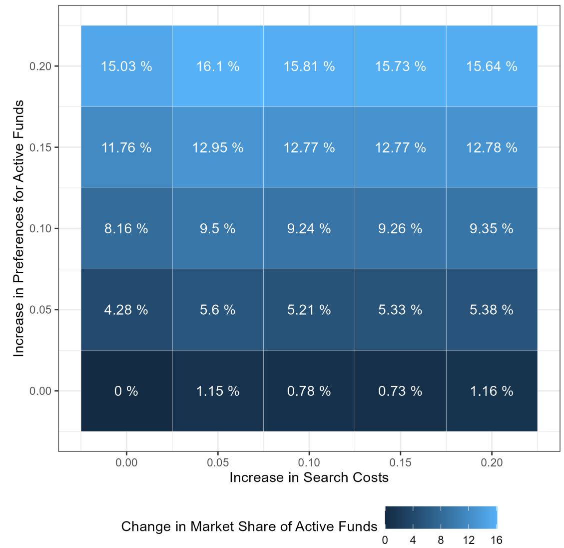
When these two counterfactual scenarios are combined, the overall impact on market size and fund pricing intricately hinges on the relative intensity of the increases in search costs and active fund preferences. Specifically, if the enhancement of preferences for active funds surpasses the rise in search costs, the market size experiences a contraction, albeit less severe than when search costs rise unopposed.

Figure 6: Counterfactual Analysis

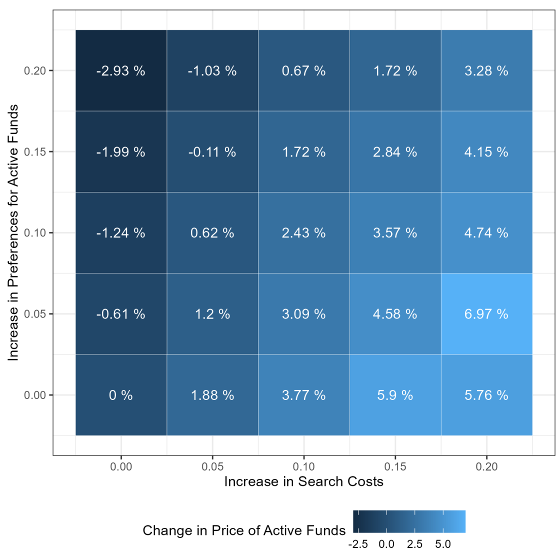
(a) Market Size



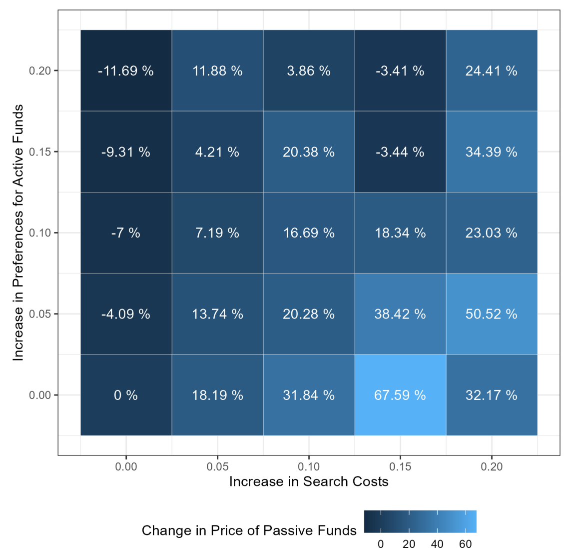
(b) Market Share of Active Funds



(c) Average Price of Active Funds



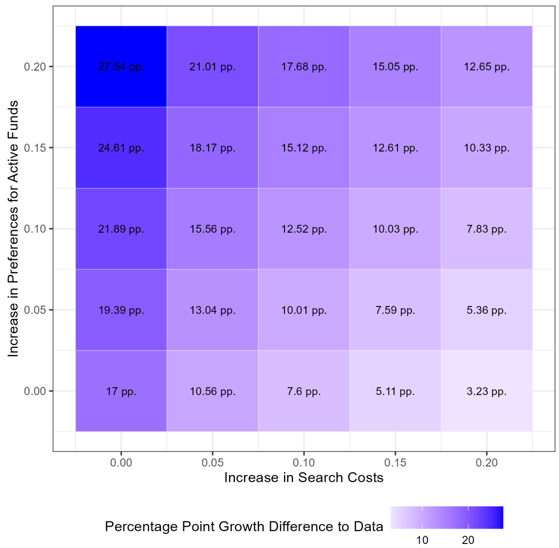
(d) Average Price of Passive Funds



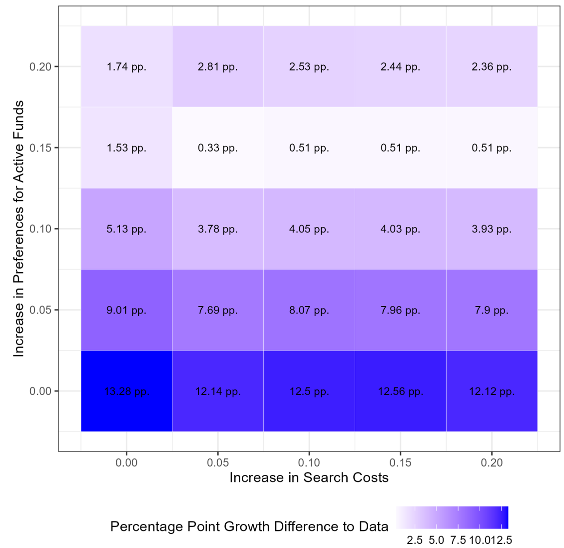
Notes: These figures present counterfactual analyses exploring the effects of increasing search costs and preferences for active mutual funds, based on 2014 data. The vertical axis in each subfigure denotes scenarios with heightened search costs, reflecting mean increases in the cost of searching for mutual funds. The horizontal axis showcases increments in the preference for active funds, adjusted such that the overall utility in the market is maintained. This adjustment ensures that while the preference for active funds is enhanced, the general inclination to invest in any fund, denoted by κ_t, remains unchanged. Each segment within the figures illustrates the outcomes of combining these two adjustments. Displayed are the results on the percentage change in overall market size, depicted as a proportion of total household assets (6a), the percentage shift in market share for active funds (6b), and the percentage variation in average pricing for both active and passive funds (6c and 6d). It is important to note that a decline in market share for active funds corresponds to an increase in market share for passive funds, highlighting the reciprocal relationship between the two.

Figure 7: Comparing the Counterfactuals to Data

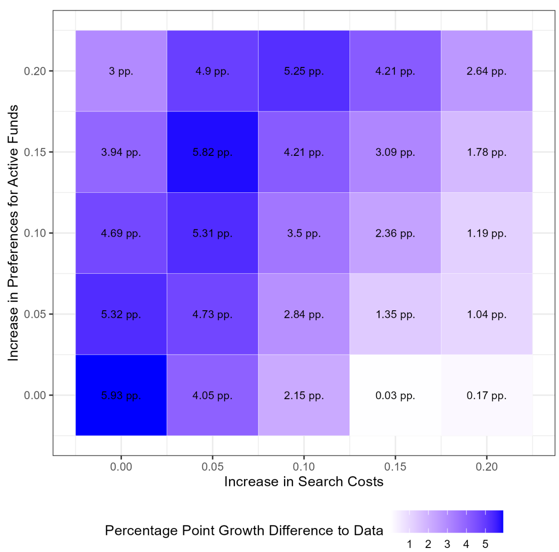
(a) Market Size



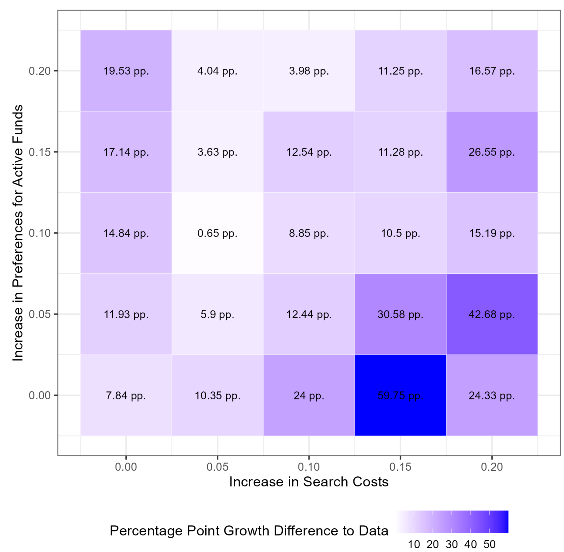
(b) Market Share of Active Funds



(c) Average Price of Active Funds



(d) Average Price of Passive Funds



Notes: The Figures show a comparison of counterfactual predictions and development in the data. The vertical axis in each subfigure represents scenarios of escalated search costs, denoting a mean increase in the effort and expense involved in fund discovery. The horizontal axis illustrates enhancements in active fund preferences, adjusted to ensure the market's average utility remains unaffected. This adjustment signifies an increased inclination towards active fund investment while maintaining the overall investment appeal, represented by κ_t. Each segment contrasts the impact of these theoretical adjustments with the observed industry transformations between 2014 and 2011, effectively analyzing the reversal in trends during this period relative to the baseline year of 2014. The displayed outcomes include percentage point discrepancies in market size (as a proportion of total household assets), active funds' market share, and the average pricing for active and passive funds (7a, 7b, 7c, and 7d). It's pertinent to note that a reduction in active funds' market share is inherently linked to an increase in passive funds' market share, highlighting the interconnected dynamics of market preferences.

8 Discussion

The development of the mutual fund market is characterized by investors shifting from active to passive funds. Despite losing market shares to passive funds prices of active funds are surprisingly constant. This article sheds light on the phenomenon. We first estimate a novel model of search that allows for preference heterogeneity. The model reveals that search costs decreased substantially over the last thirty years, and on average individuals prefer active funds. We conduct a counterfactual analysis to study the impact of search costs as well as preferences for active funds.

The structural model allows us to investigate if the development in the mutual fund market could be explained by changes in search costs or changes in preferences. We find that neither search costs nor preferences are sufficient to explain the development of the market in the past three decades. A decrease in search cost can explain an increase in market size and decreases in prices as competitive pressure increases. Further, the decrease in prices in passive funds would be higher as active funds can still rely on consumers with strong preferences. However, search cost decreases alone would not explain why passive funds increase their market share.

In comparison, lowering the preferences for active funds would have smaller effects on market size and prices but shows that the market share for active funds is decreasing. Thus, a combination of decreasing search costs and preferences for active funds could explain the market. Search costs have decreased, and the preference for active funds has become less pronounced. However, even with a decrease in preferences for active funds and lower search costs, active funds have a few customers with higher preferences. Additionally, active fund investors have, on average, higher search costs. Thus, active funds can rely on those customers and their prices are higher and decrease less than those of passive funds that experience a larger market with investors of lower search costs.

The reduction of search costs is in line with the work of Hortaçsu and Syverson (2004). However, we add to the work of Hortaçsu and Syverson (2004) by showing that the decrease in search costs can not explain the change in market shares or the price developments. We find similar to Hortaçsu and Syverson (2004) that the decrease in search costs has large effects on the number of investors entering the market. The market size increases. However, the counterfactual analysis where only search costs are reduced also reveals that a higher share of new investors would prefer to invest in active funds. Thus, we argue that part of the development is due to a change in preferences toward passive funds.

This article speaks to policymakers and regulators in the retail financial industry. Search costs are one, but not the only factor that drives consumer choice. The increased availability of stock brokers has decreased search costs for mutual funds. Nevertheless, the prices of active funds remain high. As a result, policies of reducing search costs alone do not necessarily change investing behavior. This paper shows the importance of the preferences of consumers for active funds. As shown by Gennaioli et al. (2015) one may argue that investors invest in active funds due to trust in fund managers. However, a reduction in prices for active funds relies on changes in such preferences. Regulation that solely focuses on search cost reduction may increase market size and the competitive pressure of passive mutual funds but not within the market of active funds.

To reduce search costs, several steps can be taken, such as providing clear and concise disclosure of information and developing online search tools. The US regulatory body, Securities and Exchange Commission (SEC), has taken multiple steps to reduce search costs. For example, mutual fund companies are required to provide a summary prospectus since 2009. This document contains key information about the fund and is presented in a more concise and reader-friendly format. Moreover, mutual fund companies are required to disclose a standardized measure of fund performance to facilitate comparisons across different funds (Securities and Exchange Commission, 2009). However, changing preferences or nudging investors towards less expensive and higher-performing funds is more challenging.

Two streams of literature suggest valuable insights into changing preferences or nudging investors towards less costly investments through avenues beyond just providing information. Firstly, recent studies in behavioral finance highlight the importance of peer effects in financial decisions (e.g. Bursztyn et al., 2014; Frydman, 2015; Han and Yang, 2013). As social connections and peer effects may influence financial decisions, regulators could encourage and facilitate information sharing among direct peers to nudge investors towards less expensive funds.

Secondly, evidence suggests that defaults play a crucial role in financial decisions (e.g. Carroll et al., 2009; Madrian and Shea, 2001). In situations where investment decisions are required, such as pension choices, offering low-cost defaults could be an essential step towards nudging investors towards less expensive funds.

References

Daniel Bergstresser, John MR Chalmers, and Peter Tufano. Assessing the costs and benefits of brokers in the mutual fund industry. *The Review of Financial Studies*, 22(10):4129–4156, 2008.

Steven Berry, James Levinsohn, and Ariel Pakes. Automobile prices in market equilibrium. *Econometrica: Journal of the Econometric Society*, pages 841–890, 1995.

Kenneth Burdett and Kenneth L Judd. Equilibrium price dispersion. *Econometrica: Journal of the Econometric Society*, pages 955–969, 1983.

Leonardo Bursztyn, Florian Ederer, Bruno Ferman, and Noam Yuchtman. Understanding mechanisms underlying peer effects: Evidence from a field experiment on financial decisions. *Econometrica*, 82(4):1273–1301, 2014.

Gabriel D Carroll, James J Choi, David Laibson, Brigitte C Madrian, and Andrew Metrick. Optimal defaults and active decisions. *The quarterly journal of economics*, 124(4):1639–1674, 2009.

Michael Choi, Anovia Yifan Dai, and Kyungmin Kim. Consumer search and price competition. *Econometrica*, 86(4):1257–1281, 2018.

Christopher Conlon and Jeff Gortmaker. Best practices for differentiated products demand estimation with pyblp. *The RAND Journal of Economics*, 51(4):1108–1161, 2020.

Eugene F Fama and Kenneth R French. Luck versus skill in the cross-section of mutual fund returns. *The journal of finance*, 65(5):1915–1947, 2010.

Kenneth R French. Presidential address: The cost of active investing. *The Journal of Finance*, 63 (4):1537–1573, 2008.

Cary Frydman. What drives peer effects in financial decision-making? neural and behavioral evidence. Technical report, Working paper, University of South California, 2015.

Amit Gandhi and Jean-François Houde. Measuring substitution patterns in differentiated-products industries. Technical report, National Bureau of Economic Research, 2019.

Xiaohui Gao and Miles Livingston. The components of mutual fund fees. *Financial Markets, Institutions & Instruments*, 17(3):197–223, 2008.

Nicola Gennaioli, Andrei Shleifer, and Robert Vishny. Money doctors. *The Journal of Finance*, 70(1):91–114, 2015.

Martin J Gruber. Another puzzle: The growth in activity managed mutual funds. *Journal of Finance*, 51(3):783–810, 1996.

Diane Del Guercio and Jonathan Reuter. Mutual fund performance and the incentive to generate alpha. *The Journal of Finance*, 69(4):1673–1704, 2014.

Bing Han and Liyan Yang. Social networks, information acquisition, and asset prices. *Management Science*, 59(6):1444–1457, 2013.

Elisabeth Honka. Quantifying search and switching costs in the us auto insurance industry. *The RAND Journal of Economics*, 45(4):847–884, 2014.

Ali Hortaçsu and Chad Syverson. Product differentiation, search costs, and competition in the mutual fund industry: A case study of s&p 500 index funds. *The Quarterly journal of economics*, 119(2):403–456, 2004.

Investment Company Institute. Ici factbook. Technical report, 2020. URL https://www.ici.org/pdf/2020_factbook.pdf.

Michael C Jensen. The performance of mutual funds in the period 1945-1964. *The Journal of finance*, 23(2):389–416, 1968.

Brigitte C Madrian and Dennis F Shea. The power of suggestion: Inertia in 401 (k) participation and savings behavior. *The Quarterly journal of economics*, 116(4):1149–1187, 2001.

José Luis Moraga-González, Zsolt Sándor, and Matthijs R Wildenbeest. Prices and heterogeneous search costs. *The RAND Journal of Economics*, 48(1):125–146, 2017.

José Luis Moraga-González, Zsolt Sándor, and Matthijs R Wildenbeest. Consumer search and prices in the automobile market. *The Review of Economic Studies*, 90(3):1394–1440, 2023.

Morningstar. Mutual fund database. Database, 2019. https://www.morningstar.com/.

Jennifer F Reinganum. A simple model of equilibrium price dispersion. *Journal of Political Economy*, 87(4):851–858, 1979.

Securities and Exchange Commission. Sec proposes measures to improve regulation of fund distribution fees and provide better disclosure for investors. Technical report, Securities and Exchange Commission, 2009. https://www.sec.gov/news/digest/2010/dig072110.htm.

Erik R Sirri and Peter Tufano. Costly search and mutual fund flows. *The journal of finance*, 53(5): 1589–1622, 1998.

Dale O Stahl. Oligopolistic pricing with sequential consumer search. *The American Economic Review*, pages 700–712, 1989.

The Board of Governors of the Federal Reserve System. Federal reserve economic data (fred) database. Database, Federal Reserve Bank of St. Louis, 2019. https://fred.stlouisfed.org/.

Martin L Weitzman. Optimal search for the best alternative. *Econometrica: Journal of the Econometric Society*, pages 641–654, 1979.

Asher Wolinsky. True monopolistic competition as a result of imperfect information. *The Quarterly Journal of Economics*, 101(3):493–511, 1986.

A Omitted proofs

A.1 Derivation of closed-form solution for reservation value equation

We now prove that equation (4) can be written as equation (12) as the number of products becomes large. Note that the right-hand side of equation (4) is

$$\mathbb{E}[\max\{\delta_{jt} + \mu_{ijt} + \varepsilon_{ijt} - \hat{u}_{it}, 0\}].$$

Using the law of iterated expectations, we have that this equals

$$\mathbb{E}[\mathbb{E}[\max\{\delta_{jt} + \mu_{ijt} + \varepsilon_{ijt} - \hat{u}_{it}, 0\}|\varepsilon_{ijt}]] = \mathbb{E}\left[\frac{1}{|\mathscr{J}_t|}\sum_{j\in\mathscr{J}_t}\max\{\delta_{jt} + \mu_{ijt} + \varepsilon_{ijt} - \hat{u}_{it}, 0\}\right].$$

Label the products in ascending order of utility, i.e. $\delta_{1t} + \mu_{i1t} < \delta_{2t} + \mu_{i2t} < \ldots < \delta_{Jt} + \mu_{iJt}$, where $J = |\mathscr{J}_t|$. Then the final term is equal to

$$\begin{aligned}
&\int_{\hat{u}_{it}-\delta_{1t}-\mu_{1t}}^{\infty}\left(\frac{1}{J}\sum_{j=1}^{J}(\delta_{jt} + \mu_{ijt} + \varepsilon_{ijt} - \hat{u}_{it})\right)\exp\{-\varepsilon_{ijt}\}d\varepsilon_{ijt} \\
&\qquad + \int_{\hat{u}_{it}-\delta_{2t}-\mu_{2t}}^{\hat{u}_{it}-\delta_{1t}-\mu_{1t}}\left(\frac{1}{J}\sum_{j=2}^{J}(\delta_{jt} + \mu_{ijt} + \varepsilon_{ijt} - \hat{u}_{it})\right)\exp\{-\varepsilon_{ijt}\}d\varepsilon_{ijt} \\
&\qquad + \int_{\hat{u}_{it}-\delta_{3t}-\mu_{3t}}^{\hat{u}_{it}-\delta_{3t}-\mu_{3t}}\left(\frac{1}{J}\sum_{j=3}^{J}(\delta_{jt} + \mu_{ijt} + \varepsilon_{ijt} - \hat{u}_{it})\right)\exp\{-\varepsilon_{ijt}\}d\varepsilon_{ijt} \\
&\qquad\qquad + \ldots \\
&\qquad\qquad + \int_{\hat{u}_{it}-\delta_{Jt}-\mu_{Jt}}^{\hat{u}_{it}-\delta_{J-1,t}-\mu_{J-1,t}}\frac{\delta_{Jt} + \mu_{iJt} + \varepsilon_{ijt}}{J}\exp\{-\varepsilon_{ijt}\}d\varepsilon_{ijt}.
\end{aligned}$$

Under Assumption 2, the integral in the k'th term of this sum (for $k > 1$) resolves to be

$$\frac{\exp\{-\hat{u}_{it}\}}{J}\Bigg(-\exp\{\delta_{k-1,t}+\mu_{i,k-1,t}\}\Bigg((J+1-k)(1-\delta_{k-1,t}-\mu_{i,k-1,t})+\sum_{j=k}^{J}(\delta_{jt}+\mu_{ijt})\Bigg) \\ +\exp\{\delta_{kt}+\mu_{ikt}\}\Bigg((J+1-k)(1-\delta_{kt}-\mu_{ikt})+\sum_{j=k}^{J}(\delta_{jt}+\mu_{ijt})\Bigg)\Bigg).$$

The first term of the sum is

$$\frac{\exp\{-\hat{u}_{it}\}}{J}\exp\{\delta_{1t}+\mu_{i1t}\}\Bigg(J(1-\delta_{1t}-\mu_{i1t})+\sum_{j=1}^{J}(\delta_{jt}+\mu_{ijt})\Bigg).$$

So the first and second term sum to

$$\frac{\exp\{-\hat{u}_{it}\}}{J}\Bigg(\exp\{\delta_{1t}+\mu_{i1t}\}\Bigg(1-\delta_{1t}-\mu_{i1t}+\sum_{j=1}^{J}(\delta_{jt}+\mu_{ijt})-\sum_{j=2}^{J}(\delta_{jt}+\mu_{ijt})\Bigg) \\ +\exp\{\delta_{2t}+\mu_{i2t}\}\Bigg((J-1)(1-\delta_{2t}-\mu_{i2t})+\sum_{j=2}^{J}(\delta_{jt}+\mu_{ijt})\Bigg)\Bigg) \\ =\frac{\exp\{-\hat{u}_{it}\}}{J}\Bigg(\exp\{\delta_{1t}+\mu_{i1t}\}+\exp\{\delta_{2t}+\mu_{i2t}\}\Bigg((J-1)(1-\delta_{2t}-\mu_{i2t})+\sum_{j=2}^{J}(\delta_{jt}+\mu_{ijt})\Bigg)\Bigg).$$

All subsequent terms cancel in a similar fashion. Adding up then gives that

$$\mathbb{E}[\max\{\delta_{jt}+\mu_{ijt}+\varepsilon_{ijt}-\hat{u}_{it},0\}]=\frac{\exp\{-\hat{u}_{it}\}}{|\mathscr{J}_t|}\sum_{j\in\mathscr{J}_t}\exp\{\delta_{jt}+\mu_{ijt}\}.$$

A.2 First-order conditions

We derive the first-order conditions of the profit function of firm f in market t

$$\pi_{ft}(p_t,\hat{u}_t)=\sum_{j\in\mathscr{F}_{ft}}d_j(p,\hat{u}_t)(p_{jt}-c_{jt}),$$

where p_t is a vector of all prices, $\hat{u}_t$ is the period-t distribution of reservation utilities and $\mathscr{F}_{jt}$ the set of products sold by firm f.

To derive the first-order condition, we need to derive the demand function off the equilibrium path. To do so, assume that all prices except p_i are at their equilibrium values. For convenience, we denote all variables that are held at their equilibrium values with a $*$. The demand for product j is then

$$d_j(p_{jt}, p_t^*, \hat{u}_t^*) = \frac{M_t}{\mathscr{J}_t} \int_i \frac{F^{\varepsilon}(\hat{u}_{it}^*)}{\sum_{l \in \mathscr{J}_t} \exp\{\delta_{lt}^* + \mu_{ilt}^* - \hat{u}_{it}^*\}} \exp\{\delta_{ijt} + \mu_{ijt} - \hat{u}_{it}^*\} dF^{\eta}(\eta_{it})$$

The demand for product $k \neq j$ is

$$d_k(p_{jt}, p_t^*, \hat{u}_t^*) = \frac{M_t}{\mathscr{J}_t} \int_i \frac{F^{\varepsilon}(\hat{u}_{it}^*)}{\exp\{\delta_{lt} + \mu_{ilt} - \hat{u}_{it}^*\} + \sum_{l \in \mathscr{J}_t \setminus \{j\}} \exp\{\delta_{lt}^* + \mu_{ilt}^* - \hat{u}_{it}^*\}} \exp\{\delta_{ikt}^* + \mu_{ikt}^* - \hat{u}_{it}^*\} dF^{\eta}(\eta_{it})$$

These expressions reflect that if product

For a product $j \in \mathscr{F}_{jt}$, the derivative is

$$\frac{\partial \pi_{ft}}{\partial p_{jt}} = \frac{\partial d_j}{\partial p_{jt}}(p_{jt} - c_{jt}) + d_j + \sum_{k \in \mathscr{F}_t \setminus \{j\}} \frac{\partial d_k}{\partial p_{jt}}(p_{kt} - c_{jt}), \tag{10}$$

where we suppress function arguments for legibility.

B Estimation

Our estimation procedure involves two steps. In the first step, we estimate the parameters of the utility function using standard discrete choice methods. With these parameter estimates, we move on to the second step and estimate the search costs and firms' marginal costs. This is accomplished by exploiting the consumers' reservation value equation and the firms' first-order conditions for profit maximization.

Preference parameters. To estimate preference parameters, we show an equivalence between our search model and a standard discrete choice model. To do so, consider the market shares conditional on search:

$$\sigma^I_{jt} = \int_i \frac{P(u_{ijt} > \hat{u}_{it})}{\sum_{k \in \mathscr{J}_t} P(u_{ikt} > \hat{u}_{it})} dF(\tilde{\eta}_i).$$

Here, $\tilde{\eta}_i$ is the distribution of η_i conditional on searching at least once, i.e. $\tilde{\eta}_i = \eta_i | u_{i0t} < \hat{u}_{it}$.

By making the assumption that the idiosyncratic errors are exponentially distributed with shape parameter one, we obtain the Berry et al. (1995) model. In this case,

$$P(u_{ijt} \geq \hat{u}_{it}) = \exp\{-(\hat{u}_{it} - \kappa_t - \delta_{jt} - \mu_{ijt})\},$$

where κ_i is a market-specific constant. As a result, the inside market shares can be written as

$$\sigma^I_{jt} = \int_i \frac{\exp\{\delta_{jt} + \mu_{ijt}\}}{\sum_{k \in \mathscr{J}_t} \exp\{\delta_{kt} + \mu_{ikt}\}} dF(\tilde{\eta}_i). \tag{11}$$

(The reservation utilities $\hat{u}_{it}$ cancel.) Hence, the distribution of the preference parameters $(\alpha_{it}, \beta_{it})$ can be estimated using standard methods. We use the excellent PyBLP package (Conlon and Gortmaker, 2020).

We note two differences between our setting and the standard BLP setting. First, because equation (11) does not include the outside option, an additional normalization is required. The trick we use is to write demand as a function of the differences in product characteristics with an arbitrary reference product. If we index this product with r, equation (11) becomes

$$\sigma^I_{jt} = \int_i \frac{\exp\{\delta_{jt} + \mu_{ijt} - \delta_{rt} - \mu_{irt}\}}{1 + \sum_{k \in \mathscr{J}_t \setminus \{r\}} \exp\{\delta_{kt} + \mu_{ikt} - \delta_{rt} - \mu_{irt}\}} dF(\tilde{\eta}_i).$$

To estimate the model parameters, we employ the BLP method with differenced product characteristics $X_{jt} - X_{rt}$. As the differenced variables are constant across products, we cannot include a constant term in the utility specification. Nevertheless, we allow for non-parametric shocks to the utility of the reference product.

The second difference is that there is selection based on individual heterogeneity. For example, consumers who are less price-sensitive are, all other things being equal, more likely to engage in search. Rather than modeling this selection implicitly, we make a parametric assumption about the distribution of individual heterogeneity conditional on search. In particular, we let

$$\tilde{\eta}_{it} \sim N(\lambda \sigma_{i0t}, \Sigma),$$

where σ_{i0t} is the market share of the outside option and λ and Σ are parameters to be estimated. The idea is that, because the set of searching and purchasing consumers fully overlap in our model, the number of consumers that doesn't purchase is informative on the amount of selection: the more consumers search, the less selection there will be. Hence, we model the conditional distribution as a function of the (observed) market share of the outside option.

Search and marginal costs. Under the large-market assumption, consumers that search always purchase. As a result, the fraction of consumers that does not search is equal to the market share of the outside option. From this, we can back out the average search cost in the market.

To obtain an estimating equation, we first derive a closed-form solution for the reservation utility for the case where the idiosyncratic errors ε_{ijt} follow an exponential distribution. Because the exponential distribution has positive support, a consumer will always purchase product j upon inspection when $\delta_{jt} + \mu_{ijt} \geq \hat{u}_{it}$. We rule out this case.

Assumption 2 (No certain purchases)**.** $\kappa_t + \delta_{jt} + \mu_{ijt} < \hat{u}_{it}$ *for all* i, j, t.

This assumption imposes a known upper bound on the search cost s_{it} that we impose during estimation. In the Appendix, we show that under this assumption, the equation that defines consumer i's reservation utility becomes

$$s_{it} = \frac{\exp\{-\hat{u}_{it}\}}{|\mathscr{J}_t|} \sum_{j \in \mathscr{J}_t} \exp\{\delta_{jt} + \mu_{ijt}\}. \tag{12}$$

A consumer purchases the outside option when $u_{i0t} \geq \hat{u}_{it}$, the probability of which is $\exp\{-\hat{u}_{it}\}$. Denoting by σ_{i0t} the probability that consumer i purchases the outside option, we can rewrite its reservation utility to

$$\sigma_{0it} = \frac{s_{it}}{\frac{1}{|\mathscr{J}_t|}\sum_{j\in\mathscr{J}_t}\exp\{\delta_{jt}+\mu_{ijt}\}}.$$

Taking the expectation over consumers gives that

$$\sigma_{0t} = \int_i \frac{s^c_{it}+s^m_t}{\frac{1}{|\mathscr{J}_t|}\exp\{\xi_{rt}\}\sum_{j\in\mathscr{J}_t}\exp\{\delta_{jt}+\mu_{ijt}-\xi_{rt}\}} dF(\eta_{it}), \tag{13}$$

where have substituted equation (3) for s_{it} and taken out the product quality shock of the reference product, ξ_{rt}, in the denominator. This yields a left-hand side that is directly observable in the data. The sum in the denominator of the right-hand side can be computed based on the first-step BLP estimates. Given the assumed distribution of s^c_{it}, we have two unknowns per market: the search cost shock s^m_t and the utility shock ξ_{rt}. Recall that we have estimated the shocks ξ_{jt} of the non-reference products relative to product r. Therefore, an increase in ξ_{rt} increases the utility of all products. Intuitively, a higher market share for the outside option can result from either higher search costs or less attractive products offered in the market.

To disentangle the effect of the search cost shock from the utility shock, we use the supply side of the model. The intuition is simple: a change in search costs has an effect on prices, while a utility shock that affects all products equally does not. Indeed, it can easily be derived from equation (12) that $\partial\hat{u}_{it}/\partial\kappa_t = 1$. In essence, under our large-market assumption, there is no competition with the outside option: all consumers that search at least once will purchase, while the fact that consumers have passive beliefs means that firms cannot do anything to incentive consumers to search.

While we use our large-market assumption to obtain our parameters as easily expressible functions of the data, we believe the intuition behind this identification argument is more general. Generally, in a random search model, a consumer that searches will buy the outside option when two things happen: i) the consumer has inspected all options, ii) all options are worse than the outside option. Hence, a marginal change in the value of the outside option only has a first-order effect on the price of product j to the extent that comebacks occur and that product j is the best among all products. In our model, this probability is zero, but it will be small in most random search models. Changes in search costs, on the other hand, always have an effect on prices. Hence, shocks that influence the value of the inside goods versus the outside good but not search costs can disentangle the two through their effect on prices.

To derive the first order condition, note that when the error terms are exponentially distributed, demand equation (8) becomes

$$\begin{aligned} d_j(p_j, \{\hat{u}_{it}\}) &= \frac{M_t}{|\mathscr{J}_t|} \int \frac{1-\exp\{-\hat{u}_{it}\}}{\exp\{-\hat{u}_{it}\}} \exp\{\delta_{jt} + \mu_{ijt} - \hat{u}_{it}\} dF(\eta_i) \\ &= \frac{M_t}{|\mathscr{J}_t|} \int \left(\exp\{\delta_{jt} + \mu_{ijt}\} - \exp\{\delta_{jt} + \mu_{ijt} - \hat{u}_{it}\}\right) dF(\eta_i) \end{aligned}$$

It follows that the first order condition of f's profits with respect to p_{jt}, for $j \in \mathscr{F}_{ft}$ is

$$\int \left(\exp\{\delta_{jt} + \mu_{ijt}\} - \exp\{\delta_{jt} + \mu_{ijt} - \hat{u}_{it}\}\right) \left(1 - \alpha_{it}(p_{jt} - c_{jt})\right) dF(\eta_i) = 0.$$

From equation (12), we can write $\exp\{-\hat{u_{it}}\} = |\mathscr{J}_t| s_{it} / \sum_{j \in \mathscr{J}_t} \exp\{\delta_{jt} + \mu_{ijt}\}$, so that the first order condition becomes

$$\int \left(\exp\{\delta_{jt} + \mu_{ijt}\} - |\mathscr{J}_t| s_{it} \sigma^I_{jt}\right) \left(1 - \alpha_{it}(p_{jt} - c_{jt})\right) dF(\eta_i) = 0.$$

This can be rewritten as

$$c_{jt} \int \left(\exp\{\delta_{jt} + \mu_{ijt}\} - |\mathscr{J}_t| s_{it} \sigma^I_{jt}\right) \alpha_{it} dF(\eta_i) = -\int \left(\exp\{\delta_{jt} + \mu_{ijt}\} - |\mathscr{J}_t| s_{it} \sigma^I_{jt}\right) \left(1 - \alpha_{it} p_{jt}\right) dF(\eta_i).$$

We now write

$$c_{jt} = W'_{jt} \gamma + \nu_{jt},$$

where W_{jt} contain variables that determine fund j's marginal costs, γ is a vector of parameters to be estimated and ν_{jt} is an error term. Under this parametrization, the first-order condition becomes

$$W_{jt}'\gamma + \frac{\int \left(\exp\{\delta_{jt}+\mu_{ijt}\} - |\mathscr{J}_t| s_{it}\sigma_{jt}^I\right)\left(1-\alpha_{it}p_{jt}\right) dF(\eta_i)}{\int \left(\exp\{\delta_{jt}+\mu_{ijt}\} - |\mathscr{J}_t| s_{it}\sigma_{jt}^I\right)\alpha_{it} dF(\eta_i)} + \nu_{jt} = 0.$$

The distributions of δ_{jt}, μ_{ijt} and α_{it} are estimated in the first-stage BLP step. This means that there are three types of unknowns in the first order condition: i) marginal cost parameters γ, ii) market-level search cost shocks s_t^m, and, iii) the distribution of individual search shocks s_{it}^c.

Hence, for a given distribution of s_{it}^c, γ and the all values of s_t^m can be estimated using non-linear least squares. Because there is only a single non-linear parameter per market, s_t^m, this non-linear least squares problem can be solved relatively easily.

With estimates of γ and s_t^m in hand, the market-level utility shock follows directly from equation (13):

$$\xi_{rt} = \log\left(\int \frac{s_{it}^c + s_t^m}{\frac{1}{|\mathscr{J}_t|}\sum_{j\in\mathscr{J}_t}\exp\{\delta_{jt}+\mu_{ijt}\}}\right) - \log(\sigma_{0t}).$$

What remains to be estimated is the distribution of s_{it}^c. Because search costs are likely correlated with consumer preferences, we assume the following distribution:

$$\begin{pmatrix} \eta_{it} \\ s_{it} \end{pmatrix} \sim N\left(0, \begin{pmatrix} \Sigma & \sigma_{\eta\rho} \\ \sigma_{\eta\rho} & \sigma_\rho^2 \end{pmatrix}\right).$$

. The covariance matrix of η_{it} comes from the first stage BLP estimation. Hence, the only parameters to estimate is the covariance of $\log s_{it}$ and its covariance with the elements of η_{it}.

To estimate these parameters, we construct a set of moment conditions:

$$\mathbb{E}[Z^\xi \xi_{rt}] = 0,\ \mathbb{E}[Z^s s_t] = 0.$$

We use two sets of moment conditions. For the first, we use the same moments conditions as in BLP, using instruments that are uncorrelated with product quality. The second set of moments contain instruments that are uncorrelated with search costs but correlated with the value of purchasing a mutual fund.

For the first set of instruments, we employ two types of instrument. The first is a cost shifter for the endogeneity of price. We exploit the fact that mutual funds are obligated to break down their overall fees (i.e. price) into different purposes in regulatory N-SAR filings Gao and Livingston (2008). Some of these components, like marketing fees, are best understood as the mark-up a fund chooses. However, some clearly are costs. Hence, we observe part of the marginal cost of a fund and we use this as an instrument for its fees. We use custodian fees, fees paid to hold and transfer the securities of the fund, as an instrument, as it shows more variation across funds and over time than some of the other fees given in N-SAR filings. To identify the distribution of random coefficients, we use differentiation instruments in the fashion of Gandhi and Houde (2019). The basic idea is that we take instruments measuring the degree of differentiation of a product relative to products available in the market.[8]

To identify the mean level of search costs, we need an instrument that is correlated with total investment in mutual funds but uncorrelated with search costs. We use the American business cycle as an instrument. Empirical evidence suggests that mutual fund investment is positively related to the business cycle. However, there is no reason to believe that search costs in the mutual fund market follow the same cycle. Therefore, we consider this instrument to be plausibly exogenous.

To sum up, in the second stage we use the following estimation algorithm:

1. For candidate values of $(\sigma_\rho, \sigma_{\eta\rho})$, estimate marginal costs and search cost shocks by non-linear least squares.

2. Estimate $(\sigma_\rho, \sigma_{\eta\rho})$ by minimizing the GMM objective function.

[8] In detail, we use the "quadratic" version of the instruments, so we use the sums over squared differences of products characteristics compared to rivals. See Conlon and Gortmaker (2020) for details.

C Computation of Counterfactuals

A counterfactual requires solving for

- prices p_{jt} that maximize profits given the prices of other firms and consumers' reservation values;
- reservation values $\hat{u}_{it}$ that are rational given prices.

In other words, solving for a counterfactual requires simultaneously solving firms' first order conditions (10) as well as consumers' reservation utilities (12).

To do so, we require the *unconditional* distribution of search costs and consumer preferences. (Recall that in our estimation procedure, we estimate the distributions conditional on searching at least once.) To do so, we note that by Bayes' theorem,

$$f(\alpha_i, s_i) = \frac{f(\alpha_i, s_i|\text{search})P(\text{search})}{P(\text{search}|\alpha_i, si)}.$$

Here, with some abuse of notation, $f(\alpha_i, s_i)$ is the unconditional joint density of (α_i, s_i). $f(\alpha_i, s_i|\text{search})$ is the density conditional on search, which we have estimated. $P(\text{search})$ is the unconditional (with respect to (α_i, s_i)) probability of searching, which we can observe directly in the data, i.e. it equals the market share of the outside option. $P(\text{search}|\alpha_i, si)$ is the conditional probability of searching.