Essays On Mutual Funds

Hongxun Ruan

University of Pennsylvania, ruanhx009@gmail.com

Follow this and additional works at: https://repository.upenn.edu/edissertations

Part of the Finance and Financial Management Commons

Recommended Citation
https://repository.upenn.edu/edissertations/2786

This paper is posted at ScholarlyCommons. https://repository.upenn.edu/edissertations/2786
For more information, please contact repository@pobox.upenn.edu.
Essays On Mutual Funds

Abstract
In the first chapter, "Social Capital and Innovation: Evidence from Connected Holdings", I investigate how social capital affects innovation. I measure a firm's social capital with connected holdings, which is the fraction of equity of a particular firm held by mutual funds whose managers are connected to the firm's board members through educational networks. I use plausibly exogenous variation in the size of board members' networks as an instrument for connected holdings. I find higher connected holdings lead to larger number of patents granted, more patent citations, and higher firm value created by patents. Connected holdings foster innovation by helping to reduce short-term capital market pressures and to increase management job security.

The second chapter, "Marketing Mutual Funds", co-authored with Nikolai Roussanov and Yanhao Wei, we investigate marketing and distribution expenses' impact on the allocation of capital to funds and on returns earned by mutual fund investors. We develop and estimate a structural model of costly investor search and fund competition with learning about fund skill and endogenous marketing expenditures. We find that marketing is nearly as important as performance and fees for determining fund size. Restricting the amount that funds can spend on marketing substantially improves investor welfare, as more capital is invested with passive index funds and price competition decreases fees on actively managed funds. Average alpha increases as active fund size is reduced, and the relationship between fund size and fund manager skill net of fees is closer to that implied by a frictionless model. Decreasing investor search costs would also imply a reduction in marketing expenses and management fees as well as a shift towards passive investing.

Degree Type
Dissertation

Degree Name
Doctor of Philosophy (PhD)

Graduate Group
Finance

First Advisor
Jules van Binsbergen

Keywords
Innovation, Marketing, Mutual Funds, Social Capital, Structural Estimation

Subject Categories
Finance and Financial Management

This dissertation is available at ScholarlyCommons: https://repository.upenn.edu/edissertations/2786
ESSAYS ON MUTUAL FUNDS

Hongxun Ruan

A DISSERTATION

in

Finance Department

For the Graduate Group in Managerial Science and Applied Economics

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the

Degree of Doctor of Philosophy

2018

Supervisor of Dissertation

______________________________
Jules van Binsbergen, The Nippon Life Associate Professor of Finance

Graduate Group Chairperson

______________________________
Catherine Schrand, Celia Z. Moh Professor, Professor of Accounting

Dissertation Committee

Nikolai Roussanov, Moise Y. Safra Associate Professor of Finance

David Musto, Ronald O. Perelman Professor in Finance
ACKNOWLEDGEMENT

I thank my advisors Jules van Binsbergen, Nikolai Roussanov and David Musto for their invaluable guidance. I also thank Joao Gomes for continuous support and encouragement throughout my PhD studies.

I am grateful to my friends at Wharton and Penn Econ for their help and support: Aymeric Bellon, Anna Cororaton, Amora Elsaify, Xiang Fang, Qing Gong, Deeksha Gupta, Jessica Jeffers, Mete Kilic, Jianan Liu, Yang Liu, Sebastien Plante, Lin Shen, Yanhao Wei, Ram Yamarthy and Weilong Zhang. Thank you all for friendships that will last forever!

I am extremely lucky to have Yingfu Tang, my loving wife going through the whole journey with me. I am grateful for her patience and unconditional support which makes the process enjoyable. I thank my parents-in-law for their understanding and help. And I thank Guoguo, my daughter for making all of these worth it.

Finally, I thank my parents Benqing Ruan and Chengli Wang for all their support and love! I am extremely lucky to be their son!
ABSTRACT

ESSAYS ON MUTUAL FUNDS

Hongxun Ruan
Jules van Binsbergen

In the first chapter, “Social Capital and Innovation: Evidence from Connected Holdings”, I investigate how social capital affects innovation. I measure a firm’s social capital with connected holdings, which is the fraction of equity of a particular firm held by mutual funds whose managers are connected to the firm’s board members through educational networks. I use plausibly exogenous variation in the size of board members’ networks as an instrument for connected holdings. I find higher connected holdings lead to larger number of patents granted, more patent citations, and higher firm value created by patents. Connected holdings foster innovation by helping to reduce short-term capital market pressures and to increase management job security.

The second chapter, “Marketing Mutual Funds”, co-authored with Nikolai Roussanov and Yanhao Wei, we investigate marketing and distribution expenses’ impact on the allocation of capital to funds and on returns earned by mutual fund investors. We develop and estimate a structural model of costly investor search and fund competition with learning about fund skill and endogenous marketing expenditures. We find that marketing is nearly as important as performance and fees for determining fund size. Restricting the amount that funds can spend on marketing substantially improves investor welfare, as more capital is invested with passive index funds and price competition decreases fees on actively managed funds. Average alpha increases as active fund size is reduced, and the relationship between fund size and fund manager skill net of fees is closer to that implied by a frictionless model. Decreasing investor search costs would also imply a reduction in marketing expenses and management fees as well as a shift towards passive investing.
# TABLE OF CONTENTS

ACKNOWLEDGEMENT ......................................................... ii

ABSTRACT ........................................................................ iii

LIST OF TABLES ................................................................. vi

LIST OF ILLUSTRATIONS ..................................................... vii

CHAPTER 1: Social Capital and Innovation: Evidence from Connected Holdings 1
  1.1 Introduction ............................................................. 1
  1.2 Data and Variable Constructions ................................. 7
  1.3 Empirical Strategy ................................................... 14
  1.4 Result ...................................................................... 20
  1.5 Potential Mechanisms .............................................. 31
  1.6 Conclusion ............................................................. 39

CHAPTER 2: Marketing Mutual Funds ................................. 50
  2.1 Introduction ............................................................. 50
  2.2 Model ................................................................. 57
  2.3 Data ...................................................................... 63
  2.4 Estimation ............................................................. 65
  2.5 Results ................................................................. 71
  2.6 Counterfactual Analysis ........................................... 81
  2.7 Concluding Remarks ............................................... 88

APPENDIX ................................................................. 102

BIBLIOGRAPHY ............................................................. 155
TABLE 1.1 : First Stage ......................................................... 40
TABLE 1.2 : Connected Holdings and Innovation .......................... 41
TABLE 1.3 : Disaggregating Industry: High-Tech Industry versus the Rest ... 42
TABLE 1.4 : Disaggregating Fund Type: Active versus Closet Indexers .... 43
TABLE 1.5 : Connected Holdings and R&D ................................... 44
TABLE 1.6 : Connected Holdings and Innovation Efficiency ................. 45
TABLE 1.7 : Connected Holdings and Patent Nature .......................... 45
TABLE 1.8 : Connected Holdings and Firm Growth ............................ 46
TABLE 1.9 : Trading Behavior ...................................................... 46
TABLE 1.10 : Return Behavior ..................................................... 47
TABLE 1.11 : Connected Holdings and Takeover Exposure .................... 48
TABLE 1.12 : Connected Mutual Fund Votes on Governance Issues .......... 49

TABLE 2.1 : Investor Beliefs and Manager Skill Parameters .................. 96
TABLE 2.2 : Search Model Parameters ........................................... 96
TABLE 2.3 : Change in Size when Marketing Expenses Increase by 1 bp .... 97
TABLE 2.4 : Change in Profits when Marketing Expenses Increase by 1 bp ... 98
TABLE 2.5 : Quantifying the Importance of Sampling Probability Components . 99
TABLE 2.6 : Summary of Outcomes for Current Equilibrium and No-Marketing Equilibrium ................................................................. 100
TABLE 2.7 : Summary of Outcomes for Different Search Costs ............... 101

TABLE A.1 : Variable Definitions ................................................ 108
TABLE A.2 : Days between Patent Filing Date and Grant Date ................ 109
TABLE A.3 : Number of Connections ............................................ 109
TABLE A.4 : Top 10 Most Connected University ................................ 110
<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>A.5</td>
<td>Summary Statistics</td>
</tr>
<tr>
<td>A.6</td>
<td>Connected Holdings and Innovation (Poisson Model)</td>
</tr>
<tr>
<td>A.7</td>
<td>Connected Holdings and Innovation Controling for CEO FE</td>
</tr>
<tr>
<td>A.8</td>
<td>Connected Holdings (Management Only) and Innovation</td>
</tr>
<tr>
<td>A.9</td>
<td>Connected Holdings (Broad) and Innovation</td>
</tr>
<tr>
<td>A.10</td>
<td>Connected Holdings versus Institutional Holdings</td>
</tr>
<tr>
<td>A.11</td>
<td>Citation adjustment factors</td>
</tr>
<tr>
<td>A.12</td>
<td>Daily Trading Volume and Connected Holdings</td>
</tr>
<tr>
<td>A.13</td>
<td>Summary Statistics for Termination and Connectedness</td>
</tr>
<tr>
<td>A.14</td>
<td>Relationship between Termination and Connectedness</td>
</tr>
<tr>
<td>A.15</td>
<td>Connected Mutual Fund Votes on “Say-on-Pay” Issue</td>
</tr>
<tr>
<td>A.16</td>
<td>Mutual Fund Trading and Hedge Fund Activism</td>
</tr>
<tr>
<td>A.17</td>
<td>Active Funds versus Passive Funds</td>
</tr>
<tr>
<td>A.18</td>
<td>Data Definition</td>
</tr>
<tr>
<td>A.19</td>
<td>Summary Statistics</td>
</tr>
<tr>
<td>A.20</td>
<td>Summary of Outcomes for Current Equilibrium and No-Marketing Equilibrium</td>
</tr>
<tr>
<td>A.21</td>
<td>Sensitivity of Parameter Estimates to Moments</td>
</tr>
</tbody>
</table>
LIST OF ILLUSTRATIONS

FIGURE 2.1 : Capital (mis)Allocation in Mutual Funds: Size vs. Net Skill . . . 90
FIGURE 2.2 : Capital (mis)Allocation in Mutual Funds: Net Alpha vs. Net Skill 91
FIGURE 2.3 : Marketing and Skill . . . . . . . . . . . . . . . . . . . . . . . . . . 91
FIGURE 2.4 : Price and Performance Differences: Current Equilibrium vs. No-
Marketing Equilibrium . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 92
FIGURE 2.5 : Indirect Utilities: Current Equilibrium vs. No-Marketing Equilib-
rium . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 93
FIGURE 2.6 : Heterogeneous Effect Across Investors . . . . . . . . . . . . . . . 94
FIGURE 2.7 : Capital (mis)Allocation: Counterfactual . . . . . . . . . . . . . . 95
FIGURE A.1 : Total Number of Patent Applications . . . . . . . . . . . . . . . 119
FIGURE A.2 : Growth Rate of the Total Number of Patent Applications . . . . 120
FIGURE A.3 : Total AUM of Domestic Active Funds as a Fraction of All Funds . 121
FIGURE A.4 : Histogram of Distance between Funds and Firms . . . . . . . . . 122
FIGURE A.5 : AUM and Price of Vanguard Index Funds . . . . . . . . . . . . . . 146
FIGURE A.6 : Histogram of Effective Marketing Expenses . . . . . . . . . . . . 147
FIGURE A.7 : Total Marketing . . . . . . . . . . . . . . . . . . . . . . . . . . . . 148
FIGURE A.8 : Prior Distribution of Management Skill . . . . . . . . . . . . . . 149
Net Skill . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . 149
FIGURE A.10 : Counterfactual (restricted model) . . . . . . . . . . . . . . . . . 150
CHAPTER 1 : Social Capital and Innovation: Evidence from Connected Holdings

1.1. Introduction

Economists have long been interested in the impact of social capital on economic outcomes. In the literature, there are two conflicting views: on one hand, Putnam (1993) argues that “associations instill in their members habits of cooperation, solidarity, and public-spiritedness”. The cooperative attitude and trust among group members could mitigate informational frictions and enable welfare-improving transactions.\(^1\) On the other hand, Olson (1982) argues that associational activities foster emergence of self-serving interest groups such as cartels, colluding elites and lobbies which distort markets and hurt economic growth.\(^2\) The goal of this paper is to shed new light on the relative importance of these two conflicting views. Using a measure of social connections between board members of publicly traded firms and mutual fund managers, I find that more social capital leads to more innovation at the firm level. The evidence suggests that the positive aspects of social capital dominate the potential negative ones, at least with respect to corporate innovation.

There are two major challenges associated with empirically estimating the effect of social capital on economic outcomes. The first one is measurement of social capital. The second is identification, as purely exogenous variation in social capital is rare. Previous empirical research on the question usually has focused on country-level measures of trust and has used cross-country variation to identify the impact on economic outcomes (e.g., Knack and Keefer (1997); Guiso, Sapienza, and Zingales (2009)).

To address the first challenge, I construct a firm-level measure of social capital: connected holdings, which is defined as the fraction of equity of a particular firm held by mutual funds whose managers are connected to the firm’s board members through educational networks. To address the second challenge, I exploit the plausibly exogenous variation in

\(^{1}\)See Knack and Keefer (1997), Guiso, Sapienza, and Zingales (2004), Guiso, Sapienza, and Zingales (2008), Bloom, Sadun, and Van Reenen (2012).

connected mutual fund managers' entry and exit of asset management industry and use it as an instrument for connected holdings. Since innovation is important in generating long-term economic growth, I examine the impact of connected holdings on innovation.

In a frictionless economy, social capital should not matter because financial markets can allocate the efficient amount of capital to innovative firms. In the presence of frictions such as asymmetric information, communication costs, incomplete contracting, etc., social capital could either encourage or impede firm innovation. On one hand, social capital could promote innovation by alleviating short-term capital market pressures. This might be the case if capital market investors value short-term growth in firm profits or find it hard to evaluate long-term innovative projects.\(^3\) In such an environment, management chooses to underinvest in research and development (R&D). Social capital in the form of connected holdings could alleviate capital market pressures by acting like long-term investment and supporting firms in adverse situations ("the insurance effect"). On the other hand, social capital might hurt innovation (and, more generally, firm value) because connected holdings could increase management entrenchment and weaken capital market's disciplinary effects. Classical moral hazard models inform us that under weaker monitoring, management could enjoy private benefits (e.g., through shirking or a quiet life) at the cost of firm value, which can hurt innovation (Holmström (1979), Shleifer and Vishny (1989), Bertrand and Mullainathan (2003)).

Identifying the effect of social capital on firm actions is complicated by potentially endogenous selection. Connected mutual fund managers might select firms to invest in based on the information that is observable to them, but not to the econometrician. An omitted factor could be the quality of inventors at the firm, the long-term innovation plans, or the prototypes of unpatented inventions.

I classify a firm and a mutual fund as "connected" if a fund manager and a board member attended the same university at the same time and received the same type of degree. I

\(^3\)See, for example, Shleifer and Vishny (1990), Stein (1989), Bebchuk and Stole (1993), Narayanan (1985).
address the potential selection issues by instrumenting the connected holdings with the total number of connected mutual fund managers (including those who do not hold any shares of the firm in their portfolio). More specifically, the instrument is defined as follows: in each year, for each firm in my sample, I count the number of mutual fund managers who are currently working in the asset management industry and connected to the firm. The main idea behind the instrument is that the time-series changes of the number of connected mutual fund managers affect connected holdings, but are not affected by firm innovation.

For an instrument to be valid, it needs to satisfy both the relevance condition and the exclusion condition. I formally test for the relevance condition. First-stage F-statistic estimates indicate that the instrument is highly correlated with the endogenous variable connected holdings. For the exclusion condition to hold, the instrument should not affect firm innovation through any channel other than connected holdings. After teasing out the average number of connected mutual fund managers for each board member using university fixed effects, the time-series variation in the instrument comes from the within variation of each board member’s individual connections to the mutual fund managers. And the time-series variation of each board member’s individual connections to the mutual fund managers is generated by the entry and exit of mutual fund managers, which is plausibly exogenous to firm innovation due to the fact that each connected firm on average constitutes less than 50 bps of the connected mutual fund manager’s portfolio.

Using the instrumental variable approach, I establish a causal link between connected holdings and innovation. First, I investigate the impact of connected holdings on innovation outcomes. A one standard deviation increase in connected holdings causes a 10.61% increase in the total number of patents granted, a 14.86% increase in the total number of citations, and a 2.21% increase in firm value created by patents that equates to about 23 million dollars for the average firm. Two factors affect innovation outcomes: input into innovation (Hausman, Hall, and Griliches (1984)) and innovation efficiency (Hirshleifer, Hsu, and Li (2013)). A one standard deviation increase in connected holdings translates
into a 12.61% increase in firm’s input into innovation as measured by R&D expenditure. This result suggests that more connected holdings enable firms to pursue riskier innovative projects, a finding consistent with the insurance effect of connected holdings. Relationship between connected holdings and commonly used measures of innovation efficiency, such as the ratios of either patents or citations to R&D capital, is not statistically significant. This result indicates that the negative effects induced by connected holdings are offset by positive effects, otherwise we should see a significant drop in innovation efficiency. More interestingly, innovation efficiency as measured by the ratio between the value created by patents and R&D capital is positively associated with connected holdings. This implies that connected holdings might incentivize firms to produce more valuable patents, but not more cited patents.

The impact of connected holdings on innovation is heterogeneous across firms and funds. I find that firms in industries that most rely on patents (such as pharmaceutical and IT industry), their innovations more sensitively react to the change in connected holdings. This suggests that firms in more innovative industries might be more constrained by capital market pressures. Consistent with Aghion, Van Reenen, and Zingales (2013), I find that the connected holdings coming from relatively passive funds have no impact on innovation, whereas connected holdings from truly active funds positively affect innovation. This result is intuitive: in bad situations, active funds have more degree of freedom to support the firm while passive funds need to follow some exogenously set rules which might restrain them from supporting the firm. In addition to innovation, I check how connected holdings affect other dimensions of the firm performance. I find that higher connected holdings predict higher firm growth in profits, output, and the number of employees. This finding is consistent with Lins, Servaes, and Tamayo (2017).

How do connected holdings encourage innovation? Based on previous theories, I focus on two explanations: (1) connected holdings alleviate short-term capital market pressures (Stein (1989)) and (2) connected holdings increase management’s job security (Manso (2011); Stein
Capital market pressures force firms to focus excessively on short-term earnings at the cost of long-term investment, such as innovation. I hypothesize that connected holdings could potentially reduce firm short-termism through loyal investments. Meanwhile, connected holdings could increase job security of the management by reducing takeover risk through their deterrence to corporate raiders. Consistent with both theories, I find that, first, after a firm misses their quarterly earnings target, connected mutual funds stick with the firm, whereas non-connected funds divest significantly. Second, I find that, after firms miss their quarterly earnings targets, the stock returns (as measured by one-month-ahead abnormal returns) drop less for the firms with connected holdings. Third, I find connected holdings reduce firm’s takeover exposure which increases job security of the board members. In addition, I also find that, on average, connected funds are more likely to vote against shareholder-initiated proposals on various governance issues than are non-connected funds.

My findings have two broad implications. The literature on corporate governance typically views management entrenchment as harmful to shareholder values. In this paper, I show that, when the firm wants to incentivize management to be more innovative, the firm should allow for some degree of management entrenchment. The increase in job security could increase management’s incentives to take on more risks associated with innovation. My paper also sheds new light on the role of active funds in the economy. Previous literature holds the view that active funds cannot deliver outperformance to investors, despite charging high fees. Aggregate welfare might be higher in a counterfactual world without active funds. In this paper, I show that some active funds (the connected ones) can encourage firms to be more innovative. Since innovation can promote long-run economic growth. Without those connected active funds, the economy as a whole might be worse off.

1.1.1. Related Literature

This paper links to several strands of the literature. First, it is related to papers examining the impact of social capital on economic outcomes. Putnam (1993) finds that local government performance is positively associated with people’s participation in public ac-
activities in Italy. Knack and Keefer (1997) use trust as a measure of social capital and find that social capital fosters economic growth in a cross-country regression. Similarly, La Porta et al. (1997) examine the effect of trust on the performance of large organizations in a cross-country setting and find a positive relationship. In the finance setting, Guiso, Sapienza, and Zingales (2004, 2008) document that trust affects stock market participation and international trade. In explaining the positive relationship between social capital and economic growth, Bloom, Sadun, and Van Reenen (2012) show that high social capital in an area increases decentralized decision-making within firms. To the best of my knowledge, this paper is the first to examine the role of innovation as a potential channel by which to explain how social capital affects aggregate economic growth. And I provide identification for this mechanism.

Second, my paper links to the literature on managerial short-termism. Theoretical papers argue that managers are biased toward short-term gains because of reputation concerns (Narayanan (1985), Holmström (1999)), takeover threats (Stein (1988)), and concerns about stock price (Stein (1989)). Empirical paper finds that the majority of financial executives would give up positive net present value (NPV) projects to avoid missing their quarterly earnings target (Graham, Harvey, and Rajgopal (2005)). Also, some papers call into question the existence of “short-termism” (Kaplan (2017)). I contribute to this literature by showing that social capital, as measured by connected holdings, could alleviate short-term capital market pressures, increase management job security, and lengthen firms’ planning horizon.

Third, this paper also contributes to a voluminous literature that explores the impact of financial market on corporate innovation. Bernstein (2015) investigates the impact of going public on innovation. Seru (2014) examines the effects of mergers and acquisitions on innovation. He and Tian (2013) show that financial analyst coverage causes firms to innovate less. Brav, Jiang, Ma, and Tian (2017) demonstrate how hedge fund activism reshapes innovation. My contribution to this literature is to show how the social networks
of the board members could influence innovation. Or, more broadly, how important is external investor’s trust to innovation?

Lastly, this paper links to a growing literature investigating the real impact of institutional investors on corporate policies, such as leverage (Michaely, Popadak, and Vincent (2014)), dividends (Grinstein and Michaely (2005)), R&D (Bushee (1998), Aghion, Van Reenen, and Zingales (2013)), and governance (Appel, Gormley, and Keim (2016)). The closest paper to mine is Aghion, Van Reenen, and Zingales (2013). The authors study the total impact of institutional investors on innovation outcomes. I deepen the understanding of institutional investors’ role on innovation by showing that among all institutional ownership, connected mutual fund ownership matters a lot for innovation. And surprisingly, for some innovation measures, after accounting for connected holdings, institutional ownership is not statistically significant anymore, which indicates that connected holdings might be the part of the institutional ownership that drive innovation.

The rest of the paper is organized as follows: Section 1.2 describes the data and the measures. Section 1.3 presents the identification strategy. Section 1.4 shows the results. Section 1.5 discusses the potential mechanisms. Section 1.6 concludes.

1.2. Data and Variable Constructions

I use several sources to collect data on patents, mutual fund holdings, the educational background of corporate officers and mutual fund managers, and firm-specific and fund-specific characteristics.

1.2.1. Patent Data

1.2.1.1. Innovation Measures

To measure firms’ innovative activities, I construct three measures. The first measure is the firm’s total number of patent applications filed in a given year that are eventually granted. There are two important dates for a patent, the filing date and the issue date. The filing
date issued by the USPTO on a patent application is of critical importance to an inventor or patent owner because it determines who has priority over the right to file a patent application for an invention. Once the original or amended claims of a patent application have been approved by the examiner, the patent is granted and the specific date of grant is called the issue date. Here, I use the patent’s filing year because it allows me to capture the actual time of innovation.4

However, not all patents are of equal importance. Simple patent counts cannot distinguish between groundbreaking innovation and incremental discoveries (e.g., Griliches (1990)). Hence, I construct the second measure, which is a firm’s total number of citations for patents filed in a given year. For example, suppose IBM filed for 10 patents in 1990. Then I track all the citations that those 10 patents received in subsequent years (until the end of the sample) and aggregate those citations together to obtain IBM’s total number of citations for 1990.

The above two measures mainly utilize the patent data. The last measure I used is based on Kogan, Papanikolaou, Seru, and Stoffman (2017) (KPSS, hereafter). Kogan, Papanikolaou, Seru, and Stoffman (2017) propose a new measure of the private, economic value of new innovations that is based on the stock market’s reaction to patent grants. The basic idea is to first compute the abnormal stock returns that can be attributed to patent issuance and then multiply the return with the market cap to get the value of the patent. This measure can assess the importance of each patent but also label each patent with its economic value. This measure helps me better investigating more quantitative questions related to innovation, such as how much value is created by innovative activities for a given firm. Another advantage of this measure is that, because of the forward-looking feature of the stock market’s reaction, I don’t need a long period of time to accumulate citations to access the importance of the patent. (For more details on how Kogan, Papanikolaou, Seru, and Stoffman (2017) construct their measure, please refer to their paper.)

4 For the reasons to use the filing date, see Griliches, Pakes, and Hall (1986).
1.2.1.2. Patent Data Source

I downloaded the patent data from Professor Noah Stoffman’s Web site.\(^5\) A detailed description of this dataset can be found in Kogan, Papanikolaou, Seru, and Stoffman (2017). Kogan, Papanikolaou, Seru, and Stoffman (2017) begin with all patents downloaded from Google Patents. They matched patents to Center for Research in Security Prices (CRSP) firms by the assignee’s name. In their final dataset, they have 1,928,123 matched patents, of which 523,301 (27\%) are new compared with the commonly used NBER patent dataset (see Hall, Jaffe, and Trajtenberg (2001) for more details).

The patent data contain truncation problems that arises because patents appear in the database only after they are granted. There is a lag between the patent filing date and the patent issuing date. We can see from Table A.2 that the mean lag days in the 1980s and in the 1990s is approximately 2 years. But it significantly increases to more than 3 years in the 2000s with the 99th percentile reaching almost 8 years. In Figure A.1, I plot the total number of patent applications that are eventually granted by year. We can see a significant drop starting from year 2003. This drop is due to the lag days: a lot of the patents applied for after 2003 are still under review and had not been granted by 2010, which is the last year of the patent data. To deal with this problem, I choose 2003 as the last year of my sample. This will guarantee that 99\% of patents applied for before 2003 (including 2003) have been issued before 2010. I also check the growth rate of the total number of patent applications, and find that it is relatively stable from 1980 to 2003. Starting from 2003, it significantly decreases. For robustness, in the result section, I also rerun all the exercises with 2004 or 2005 as the last year of the sample. All the results are quantitatively the same. The truncation problem for the patent citations arises as patents tend to receive citations over an extended period of time (e.g., 20 years), but I only observe the citations up to 2010. Then patents issued near the end of the dataset have no time to accumulate citations. To deal with this problem, I follow Hall, Jaffe, and Trajtenberg\(^5\)https://iu.app.box.com/v/patents
(2001) and Hall, Jaffe, and Trajtenberg (2005). I estimate the citation-lag distribution for different technology categories. For patents that haven’t received 20 years of citations, I use a citation-lag distribution to infer the total number of citations that those patents should receive. Other methods can be used to address the citations truncation problem. For example, one can only count the total number of citations a patent received in the first N (N = 0, 1, 2, 3) years after it is granted. For robustness, I also check this measure, and all the results are qualitatively the same.

Kogan, Papanikolaou, Seru, and Stoffman (2017) measure of the value of patent is based on the stock market’s reaction on the patent issue date. As discussed before, the patent filing date is more informative about the actual time the innovation occurred. So I move the KPSS measure to the patent application year.

Following the innovation literature, I set the patent and citation counts to zero for firms without available patent or citation information from the patent database. The distribution for both patent applications and patent citations is right skewed, with its median at zero. To deal with skewness, first I winsorize those variables at the 99th percentile, and, second, I take the natural logarithm of both patent applications (\(LnPatApp\)) and patent citations (\(LnPatCite\)). This creates another problem: a lot of the observations with zero patents become missing values after taking the natural logarithm. I take two approaches: (1) I drop the observations with missing values (Aghion, Van Reenen, and Zingales (2013) also takes this approach), and (2) to avoid losing observations, I add one to the actual values before calculating the natural logarithm. I denote those cases as (\(Ln\hat{PatApp}\)) and (\(Ln\hat{PatCite}\)). I also apply the above procedures to the KPSS measure of the patent value. I performed a similar transformation, which I denote as \(LnKPSS\) and \(Ln\hat{KPSS}\).

\footnote{For more details, see Appendix A1.6, where I provide the citation-lag distribution and adjustment methods.}
1.2.2. Mutual Fund Holdings and Fund Manager Data

My data on mutual fund holdings come from the Thomson Reuters CDA/Spectrum S12 database, which includes all registered mutual funds filing with the SEC. I only include common stock holdings of mutual funds (i.e., a stock with the share code 10 or 11 in CRSP).

I obtain portfolio managers’ biographical information from Morningstar, Inc. For each mutual fund manager, Morningstar provides the manager’s name, all college and graduate degrees he or she received, the year in which the degrees were granted, and the granting institution. Morningstar also provides the employment history for each mutual fund manager, including the fund name, starting date, and end date. I first merge Morningstar dataset with CRSP mutual fund dataset, then I use MFLINKS data link provided by Wharton Research Data Services to merge it with Thomson Reuters fund holdings dataset (see Wermers (2000) for details on how to merge these two databases). My final mutual fund sample includes survivorship-bias-free data on holdings and biographical information for 3,094 U.S. mutual funds and 5,369 mutual fund managers between January 1980 and December 2003.\(^7\)

\[\text{INSERT FIGURE A.3 HERE}\]

1.2.3. Company Officers Data

The senior biographical information of company officers (defined as Chief Executive Officer [CEO], Chief Financial Officer [CFO], Chief Technological Officer [CTO], Chief Operating Officer [COO], and Chairman) and board of directors was obtained from BoardEx of Management Diagnostic Limited, a private research company specializing in collecting and disseminating social network data on company officials in U.S. and European public and private companies and other types of organizations (e.g., charity). For each senior company

\[^7\]Notice that I keep all the matched funds instead of only active domestic equity mutual funds because I assume all kinds of funds’ holdings could potentially affect firm’s innovation. Out of those 3,094 mutual funds, 2,037 are domestic, well-diversified equity mutual funds. The domestic, well-diversified equity mutual funds compose about 90% of all the Asset Under Management (AUM) in the final dataset. The time series of total AUM of domestic active funds as a fraction of all funds can be found in Figure A.3.
officer and board of directors, the Boardex database provides all the college and graduate
degrees received, the year in which the degrees were granted, and the granting institution.
Boardex also provides the employment history for each company official. This information
includes his or her current and past roles with a start date and end date, as well as a dummy
indicating whether the individual serves (served) on the board of directors in the current
(past) employment position. Given this paper’s focus, I restrict the sample to U.S. public
firms.

1.2.4. Connections and Connected Holdings

Following Cohen, Frazzini, and Malloy (2008), I define the social networks over educational
institutions. I group the degrees for mutual fund managers, senior company officers, and
board of directors into six categories: (1) undergraduate, (2) Master of Business Adminis-
tration (MBA), (3) general graduate (MA or MS), (4) doctorate (PhD), (5) law school (JD),
and (6) medical school (MD). The broad connection dummy equals 1 if the fund manager
and a senior officer of the firm and/or a board member of the firm attended the same uni-
versity. The narrow connection dummy equals 1 if the fund manager and a senior officer
of the firm and/or a board member of the firm attended the same university at the same
time and received the same type of degree (from the above-defined six groups of degrees).
In terms of connectedness, narrow connections should dominate broad connections.

After I define the connection, connected holdings for a firm in a given year is defined by the
following formula:

\[
\text{ConHold}_{i,t}^k = \frac{\sum_{j=1}^{N_t} \text{CONNECTED}_{i,j,t}^k \cdot S_{i,j,t}}{\bar{S}_{i,t}} \quad k = \text{narrow, broad},
\]

where \(\text{CONNECTED}_{i,j,t}^k\) is a dummy variable that indicates whether fund manager \(j\) is
connected to firm \(i\) through a type \(k\) connection in year \(t\). \(S_{i,j,t}\) is the shares held by the
mutual fund manager \(j\) in firm \(i\) at time \(t\). \(\bar{S}_{i,t}\) is the total shares outstanding for firm \(i\)

\[^8\text{I correct the inconsistencies in the educational institution names. The appendix provides the details.}\]
at time $t$. $N_t$ is the total number of mutual fund managers at time $t$. $\text{ConHold}_{t}^{k}$ measures the fraction of a firm’s equity that is held by connected mutual fund managers.

1.2.5. Other Data and Control Variables

Following the literature, I control for a set of variables that may affect a firm’s innovation. To calculate the control variables, I collect financial statement items from Compustat, institutional holdings data from Thomson CDA/Spectrum database (Form 13F), stock price information from CRSP, and institutional investor classification data from Brian Bushee’s Web site. The control variables include firm age, firm sale, ROA, asset tangibility, R&D stock, R&D investment intensity, leverage, Tobin’s Q, investment intensity, institutional investor’s ownership, dedicated institutional holdings, transitory institutional holdings, and quasi-indexed institutional holdings. I provide detailed variable definitions in Table A.1.

1.2.6. Summary Statistics

Table A.3 reports the number of connections for the top 15 educational institutes according to the definition of CONNECTED(broad) and CONNECTED(narrow), respectively. Harvard University and the University of Pennsylvania have the most connections under both definitions.

To mitigate the impact of outliers, I winsorize all variables at the 1st and 99th percentiles. Table A.5, Panel A, reports the summary statistics for the full sample. Panel B separately shows statistics for those with positive CONNECTED(narrow) holdings and those with no CONNECTED(narrow) holdings. A salient feature in the data is that the firm-year observations with positive CONNECTED(narrow) holdings have significantly higher

---

9http://acct3.wharton.upenn.edu/faculty/bushee
innovation outcomes as measured by the total number of patent applications (which were eventually approved), the total patent citations, and the KPSS value. Firms with positive CONNECTED(narrow) holdings also have a higher ROA, higher R&D stocks, and a higher Tobin’s Q. But across the measures regarding routine investment, such as PPE or CAPX, the firms with positive CONNECTED(narrow) holdings have lower values.

The previous literature (Bushee (1998); Aghion, Van Reenen, and Zingales (2013)) documents that among all different types of institutional investors, dedicated investors are important in terms of encouraging R&D investment and fostering innovation. The ratio between ConHold(broad) and dedicated institutional holdings in the sub-sample with positive CONNECTED(narrow) holdings is about 62.8%. And the ratio between ConHold(narrow) and dedicated institutional holdings is about 6.7%. So connected holdings constitute a sizable fraction of the dedicated holdings. In the following sections, without special notice, all the connected holdings are ConHold(narrow).

1.3. Empirical Strategy

1.3.1. Specification

To explore the relationship between connected holdings and innovation, my baseline specification is as follows:

\[ y_{i,t} = \alpha + \beta \text{ConHold}_{i,t} + X'_{i,t} \gamma + \eta_i + \mu_t + U_{i,j,t} + \epsilon_{i,t}, \]  

(1.2)

where the indices \( i, j, \) and \( t \) correspond to the firm, the university, and the year, respectively. \( y_{i,t} \) is various measures of innovation for firm \( i \) in year \( t \). ConHold\(_{i,t} \) is the fraction of the firm’s equity held by connected mutual funds as defined in equation (1.1). \( X'_{i,t} \) is a vector of firm and industry control variables that may affect a firm’s innovation output. \( \eta_i \) is firm
fixed effects. $\mu_t$ is the time fixed effects, and $U_{i,j,t}$ is a vector of dummies for each university in my sample. It equals 1 if one of the board members from firm $i$ graduated from university $j$ at time $t$. I cluster standard errors at the firm level. The appendix examines the other specifications, such as the Poisson model.

1.3.2. Identification

To establish the causal link between connected holdings (the proxy for social capital of the firm) and innovation is challenging. There are two selection issues. The first one is the endogenously chosen level of connected holdings. Connected mutual funds might select firms in which to invest based on the information observable to them, but not to me, the econometrician. An omitted factor could be the quality of inventors at the firm, the long-term innovation plan, or the prototype of some unpatented inventions. Such factors clearly affect innovation and are correlated with connected holdings. Apparently, constructing direct measures of those factors is difficult. But without the appropriate control, my results will be biased.

The second selection issue is that the educational network could be endogenous. People often self-select into groups (Manski (1993); Angrist and Pischke (2008)). In my setting, mutual fund managers and corporate officers might choose to attend a university based on characteristics that are unobservable to the econometrician. For example, the average risk appetite of Harvard graduates might be different from that of Stanford graduates. And those unobserved characteristics might affect the connected holdings and the innovation and, in doing so, would cause omitted variable bias.

To tackle the first selection issue, I use an instrumental variable (IV) method. To address the second issue, I include university fixed effects as a control. I will first discuss the IV method, and in the exclusion section of my IV method, I discuss the importance of using university fixed effects and how they address the second selection issue.

My IV is the total number of connected mutual fund managers to the firm (including mutual
funds that do not hold shares in the firm). More specifically, it is defined as follows: in each year, for each firm in my sample, I count the number of mutual fund managers who are “active” (i.e., working) and connected to the firm through an educational link. In the following two subsections, I discuss the necessary assumptions that need to hold for the instrument to be valid.

1.3.2.1. Relevance Condition

For the instrument to be valid, it must strongly affect connected holdings. This point has been partially established in Cohen, Frazzini, and Malloy (2008). There, the authors show that mutual fund managers place larger bets on connected firms. Everything else equal, a firm with more connected mutual fund managers should have higher connected holdings. To formally test the relevance condition, I estimate the following first-stage regression:

\[
ConHold_{i,t} = \delta z_{i,t} + X_{i,t}'\gamma + \eta_i + \mu_t + U_{i,j,t} + \epsilon_{i,t}, \tag{1.3}
\]

where \(z_{i,t}\) stands for the number of mutual fund managers connected to firm \(i\) at time \(t\). For the IV to be valid, we need \(\delta \neq 0\). In Table 1.1, I show the first-stage results. I find that the coefficient of the number of connected mutual fund managers is positive and significant at the 1% level. The F-statistic equals 48.44 and exceeds the threshold of \(F = 10\), suggesting that the instrument is strong (Stock and Yogo (2005)).

Using results from column 2, I show that an increase of one standard deviation in the number of connected mutual fund managers translates into an about 0.28-standard-deviation increase in CONNECTED(narrow) holdings. For robustness, I also check the first-stage regression for CONNECTED(broad). The results are similar.

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 HERE]}\]

\[\text{[INSERT TABLE 1.1 Here}
1.3.2.2. Exclusion Condition

The instrument not only needs to affect connected holdings, but importantly, it must satisfy the exclusion restriction. That is, it should be uncorrelated with the residual in equation (1.2). In other words, the instrument should influence the outcomes of firm-level innovation only through its effect on connected holdings. To validate the exclusion condition, I check the following dimensions:

Endogenous Choice of University Does Not Affect Time-Series Change in IV

The network used in this paper is based on the institutions that corporate officers and mutual fund managers both attended. There is a concern that agents (CEO or mutual fund managers) endogenously choose to attend a university based on unobserved characteristics. And those unobserved characteristics will be correlated with my instrumental variables, subsequently biasing my results (omitted variable bias). In this section, I analyze this possibility and show how to use university fixed effects to control for unobserved characteristics.

In my setting, my IV is the total number of mutual fund managers who are connected to the firm’s board members. A connection is defined as people who attend the same university at the same time to obtain the same type of degree. Along three dimensions that determine connections, university choice is the one most likely to be endogenous.\textsuperscript{11} From now on, I will focus on the endogeneity problem induced by the agent’s choice of the university.

Assume the true data-generating process for firm-level innovation is as follows:

\[
y_{i,t} = \alpha + \beta \text{ConHold}_{i,t} + X'_{i,t} \gamma + \eta_i + \mu_t + u_{i,t} + \lambda_{i,t} + \epsilon_{i,t},
\]

where all the other variables share the same definition used in Section 1.3.1. \(\lambda_{i,t}\) is a vector

\textsuperscript{11}I admit that the choice to attend university in a specific year also might be endogenous. But the time fixed effects should be able to control for this concern.
that includes all the individual fixed effects for the board members for firm \(i\) at time \(t\). In the simplest case, suppose there is only one CEO on the board. Then \(\lambda_i,t\) is equal to the CEO’s individual fixed effects. Notice that if, throughout the sample, the firm has only that CEO, then the firm fixed effects and \(\lambda_i,t\) cannot be separately identified, because there is no time variation in either variable at the firm level. But if the firm replaces their CEO at some time, then \(\lambda_i,t\) captures the CEO’s fixed effects at the corresponding periods. We can generate it to multiple person’s case. \(\epsilon_i,t\) is an error term that is uncorrelated with any regressors. I denote \(\lambda_i,t + \epsilon_i,t\) as the composite error \(u_i,t\) that could be correlated with my IV.

The instrumental variable I used is the total number of mutual fund managers who are connected to firm \(i\) at time \(t\). To make the IV invalid, we should have

\[
E[z_{i,t} \cdot u_{i,t}] = E[z_{i,t} \cdot (\lambda_{i,t} + \epsilon_{i,t})] \\
= E[z_{i,t} \cdot \lambda_{i,t}] \\
\neq 0,
\]

where the second equality comes from the fact that \(\epsilon_{i,t}\) is an error term that is uncorrelated with any regressors. Because of the way \(z_{i,t}\) is constructed, the only possible way \(\lambda_{i,t}\) is correlated with \(z_{i,t}\) must be the endogenous choice of universities of the board members based on \(\lambda_{i,t}\). But \(\lambda_{i,t}\), at most, should be correlated with \(z_{i,t}\), which is the sum of the average network size for each network of the board member. In other words, the manager’s individual fixed effects might affect his or her choice to attend Harvard University versus Stanford University. This could affect his or her average network size. But these individual fixed effects should be uncorrelated with the time-series variation of the network size. In other words, the CEO’s individual fixed effects could not affect how many mutual fund managers are currently working in the mutual fund industry at any period in time.
I should find that \((z_{i,t} - \bar{z}_{i,t})\) is uncorrelated with \(\lambda_{i,t}\). The proof is as follows: given \(E[(z_{i,t} - \bar{z}_{i,t})|\lambda_{i,t}] = 0\) because \((\bar{z}_{i,t})\) is the mean of \(z_{i,t}\).

\[
E[(z_{i,t} - \bar{z}_{i,t})|\lambda_{i,t}] = E[E[(z_{i,t} - \bar{z}_{i,t}) \cdot \lambda_{i,t}|\lambda_{i,t}]] = E[\lambda_{i,t} E[(z_{i,t} - \bar{z}_{i,t}) \cdot \lambda_{i,t}]] = E[\lambda_{i,t} \cdot 0] = 0
\]

This shows that the demeaned instrument is uncorrelated with the individual unobserved characteristics. The university fixed effects could capture this demeaning effect. I create a full set of dummies for each university and set it to 1 if any company officer receives a degree from this university.

**Board Members Cannot Affect Mutual Fund Manager’s Career**  In the last section, I argue that the board members’ individual fixed effects will not affect the time-series change in the number of connected mutual fund managers after controlling for the university fixed effects. But if board members could affect the career path of the mutual fund manager, then this point could invalidate my IV. For example, board members could lengthen the tenure of connected mutual fund managers by feeding them insider information. I check this story by investigating the relationship between fund manager’s termination and their connectedness. In the appendix, I provide the results. There is no statistically significant relationship between fund manager’s connectedness (defined as the number of board member connections) and their termination (defined as not working in the current fund in the next year). The potential reason is that, usually, fund managers don’t hold a lot of equity from connected firms because of the diversification restriction imposed by the fund family. So even if they could profit from connected holdings, those outperformances only marginally contribute to their total performances.
**Dynamic Formation of Network Might Not Exist**  
Agents might form social ties in anticipation of future economic benefits (Manski (1993)). If the links are dynamically formed (Jackson (2010)), then there could be an endogeneity problem. For example, assume the mutual fund manager A wants to know more information about firm F. And fund manager A endogenously formed a link with firm F through CEO B. If this sort of thing happens in the data, then the IV (the number of connected mutual fund managers) could be possibly driven by the profitability of the firm, a finding that could be correlated with both the IV and the innovation. But according to my network definition, this scenario is not possible because whether A and B attended the same school for the same degree at the same year is determined many years in the past. To investigate this point, I check the age distribution for the company officers and the mutual fund managers. In my data, the median age for the fund manager and company officers is 45 and 54, respectively. Most likely, they have passed the age for education. This evidence rules out the story that maybe a mutual fund manager or a company officer would endogenously choose their education to form the link.

[INSERT TABLE A.5 HERE]

1.4. Result

In this section, I examine the impacts of connected holdings on firm innovation using the instrumental variable approach described in Section 1.3.2. Throughout all specifications, I control for firm, year, and university fixed effects. I also control for the following variables: log sales, firm age, asset tangibility, leverage, Tobin’s Q, ROA, and investment rate. Table A.1 provides the variable definitions.

1.4.1. Innovation Output

In this section, I examine the relationship between connected holdings and various measures of innovation outcomes. In brief, I use three sets of dependent variables to measure the outcome of innovation: (1) the total granted patents for a firm in a given year; (2) the total
number of citations received by those patents; and (3) the value created by those patents for the firm as measured by $KPSS$.\footnote{For detailed variable definitions, see Table A.1.}

Table 1.2 reports the results. I find a statistically significant positive relationship between connected holdings and various measures of innovation output. With more connected holdings, firms produce more patents and these patents receive more citations. Meanwhile, more firm value is created by these patents. In terms of economic significance, a one-standard-deviation increase in connected holdings causes a 10.61% ($= (42.46 \times 0.0025 \times 100)\%$) increase in the number of patents generated by the firm, a 14.86% ($= (59.47 \times 0.0025 \times 100)\%$) increase in the number of citations, and a 2.21% increase in the firm value created by the patents. To put those results into perspective, a 10.61% increase in the number of patents for the average firm is equal to 6 more patents. And a 2.21% increase in firm value for the average firm is equal to 23 millions dollars ($= 2.21\% \times 1,089$).

\begin{table}[h]
\centering
\caption{Innovative Industries versus the Rest}
\begin{tabular}{ll}
\hline
Industry & Description \\
\hline
Business Equipment & Computers, Software, and Electronic Equipment. \\
Health care & Industry traditionally puts more emphasis into inventions (think about all the new drugs). \\
Low innovation & Remaining industries.
\hline
\end{tabular}
\end{table}

Innovations do not equally spread across industries.\footnote{According to Kogan, Papanikolaou, Seru, and Stoffman (2017), “developments in computing and telecommunication have brought about the latest wave of technological progress in the 1990s and 2000s, which coincides with the high values of our measure. In particular, it is argued that this is a period when innovations in telecommunications and computer networking spawned a vast computer hardware and software industry and revolutionized the way many industries operate. We find that firms that are main contributors to our measure belong to these sectors with firms such as Sun Microsystems, Oracle, EMC, Dell, Intel, IBM, AT&T, Cisco, Microsoft and Apple being the leaders of the pack.”} Connected holdings could potentially have heterogeneous impacts on firms in different industries. I split my sample into two groups by innovativeness. To begin with, I use Fama-French 12 industry categorization. For the high innovation group, I pick the Business Equipment industry, which covers Computers, Software, and Electronic Equipment. I also pick the health care industry because this industry traditionally puts more emphasis into inventions (think about all the new drugs). I put the remaining industries into the low innovation group. I find that for the high innovation group, the impact of connected holdings is much more pronounced. For patent
applications, the coefficient doubled, and for patent citations, the coefficient tripled. This means that the connected holdings are much more important for the innovation process in the industry in which innovation is important. One way to interpret this result is that the firms in the high innovation group might be more constrained. When they have more support, they are more free to innovate. But for the firms in the low innovation group, the incentive to innovate is low. So even if the firms have more connected holdings, the production of innovation still could be low. In an untabulated table, I repeat this above analysis using the Fama-French 30-industry portfolio, and the results are quantitatively similar.

[INSERT TABLE 1.3 HERE]

1.4.1.2. Active Funds versus Closet Indexers

In the previous sections, I examine the impact of total connected holdings on innovation. However, there exists significant heterogeneity across the fund types, that is, active funds versus passive funds. Active funds can influence firms’ behaviors through their actions in the public equity market, for instance, investment and divestment (Edmans (2009)). Passive funds don’t have the divest option. They mainly influence firms through the “voice” channel (Levit and Malenko (2011), Appel, Gormley, and Keim (2016)).

In terms of passive funds, there are mainly two types: (1) “true” passive funds (their goals are to replicate some existing indexes, e.g., S&P 500 Index Funds.) and (2) closet indexers (they don’t claim they are index funds, but their investment styles are similar to those of index funds). I first check the presence of index funds in my sample by mainly focusing on the number and total asset management. Table A.17 in the appendix provides the results. Because MFLINKS from WRDS mainly focus on matching active funds between CRSP and the Thomson Reuters dataset, there are very few index funds in my sample. In terms of fund numbers, the time-series average of passive funds is below 3%. In terms of total AUM, the time-series average of passive funds is below 2%. So within passive funds, I focus on
Currently, there is no universal way to detect closet indexers. Cremers and Petajisto (2009) argue that a fund with active shares between 20% and 60% are more likely to be a closet indexer.\textsuperscript{14} I choose the cut-off as 50%; that is, I label the funds with active shares lower than 50% as closet indexers. The time-series average of the AUM of closet indexers versus the total AUM in the data is 7.27%.

I redo the analysis regarding innovation by including both active funds’ connected holdings and closet indexers’ connected holdings:

\[ y_{i,t} = \alpha + \beta_1 \text{ConHold}^a_{i,t} + \beta_2 \text{ConHold}^c_{i,t} + X_{i,t}'\gamma + \eta_i + \mu_t + U_{i,j,t} + \epsilon_{i,t}, \]

where \( \text{ConHold}^a_{i,t} \) is the connected holdings for firm \( i \) in year \( t \) from active funds, and \( \text{ConHold}^c_{i,t} \) is the connected holdings from closet indexers. All the other variable definitions are similar to those in the main specification in equation (1.2). The instrumental variable for \( \text{ConHold}^a_{i,t} \) is the connected mutual fund managers for firm \( i \) in year \( t \) from active funds, and the instrumental variable for \( \text{ConHold}^c_{i,t} \) is the connected mutual fund managers for firm \( i \) in year \( t \) from closet indexers.

Table 1.4 reports the results. I find that the impact of connected holdings mainly comes from active funds instead of closet indexers. This result is similar to that of Aghion, Van Reenen, and Zingales (2013), who find quasi-indexed institutional investors have no impact on firms innovation. This result indicates that, quantitatively, active funds exert more impacts on firm’s innovation policy than passive funds.

\[ \text{[INSERT TABLE 1.4 HERE]} \]

\textsuperscript{14}Active shares are the shares of portfolio holdings of a fund that differ from benchmark index holdings.
1.4.2. Innovation Input

Potentially, there could be two factors contributing to the increase in innovation outcomes following an increase in connected holdings. Connected holdings cause (1) an increase in innovation input (i.e., R&D expenditure increases) and (2) an increase in innovation efficiency (i.e., patents per R&D dollar). There are three sets of combined reasons to explain the increases in innovation outcomes: (1) an increase in input, but a decrease or no effect in efficiency; (2) an increase in efficiency, but a decrease or no effect in input; and (3) both an increase in efficiency and in input.

It has long been recognized that input into R&D matters for the production of patents (Hausman, Hall, and Griliches (1984)). Because of the skewness of R&D, I take the natural logarithm of R&D expenditure. There are two problems with the R&D data: (1) because of firms’ incentives to disguise their behaviors, a lot of firms choose not to report R&D. In Compustat, they are missing numbers. I drop these observations. And (2) for some observations, the value of R&D is zero. After calculating the natural logarithm, these values become missing values. I take two approaches: (1) I drop the observations with missing values (Aghion, Van Reenen, and Zingales (2013) also takes this approach), and (2) to avoid losing observations, I add one to the actual values before calculating the natural logarithm. I denote those cases as \( \text{Ln}\widehat{R&D} \).

In Table 1.5, I report the results. In columns (1) and (3), I report the baseline results. I find a statistically significant positive relation between connected holdings and firm’s investment in R&D. In columns (2) and (4), after controlling for the university fixed effects, the coefficients are still significant and only marginally decrease in magnitude. The results are robust to using either \( \text{LnR&D} \) or \( \text{Ln}\widehat{R&D} \). Since I use a log-level specification, the interpretation of the coefficient is semi-elasticity. A one-standard-deviation increase in connected holdings leads to a 12.61% \( (= (50.46 \times 0.0025 \times 100)\%) \) increase in firms’ R&D expenditure. This

\[^{15}\text{In an unreported analysis, I replace all the missing values for R&D with zero, since my sample is post-FASB 1975 R&D reporting requirement. The results are quantitatively similar.}\]
increase is economically sizable. To put the results in context, the average annual growth rate of R&D in my sample is 21.37%. So the increase in R&D for the firm that experiences a one-standard-deviation increase in connected holdings is about 59% of it.

To explore an interesting question about how persistent these impacts are, I relate R&D expenditure in the next 2 years (R&D in year \( t + 2 \), not the cumulative R&D from year \( t \) to year \( t + 2 \)) to connected holdings. And I find a positive significant result. This means that the connected holdings have a long-term impact on the firm’s R&D policy. And those long-term impacts align with the long-term project feature of the innovation process, which could lead to more innovation.

1.4.3. Innovation Efficiency

In this section, I check for innovation efficiency. The null hypothesis is that connected holdings have no impact on innovation efficiency. There are two alternative hypotheses.

Hypothesis A1: Connected holdings encourage management to take on more risk. This does not mean that management takes on the risk in a discrete way. They could invest in a lot of projects, both promising and unpromising ones. Doing so leads to a decrease in innovation efficiency.

Hypothesis A2: Connected mutual fund managers provide useful information to management regarding innovation. Doing so could induce management to pick the right projects. In this case, innovation efficiency increases.

I construct three proxies for innovation efficiency: The first measure is the ratio between total patents and R&D expenditures. The second measure is the ratio between citations and R&D expenditures. The third measure is the ratio between KPSS and R&D expenditures.

Following Hirshleifer, Hsu, and Li (2013), I define innovation efficiency (IE) in the following way:
\[ IE_{i,t} = \frac{Innovation_{i,t}}{(R&D_{i,t} + 0.8 \times R&D_{i,t-1} + 0.6 \times R&D_{i,t-2} + 0.4 \times R&D_{i,t-3} + 0.2 \times R&D_{i,t-4})}, \]

where the numerator is the total number of patents, the total number of citations, and the value created by patents (KPSS) for firm \( i \) in year \( t \). The denominator is R&D capital, which is the 5-year cumulative R&D expenses, assuming an annual depreciation rate of 20% like in Chan, Lakonishok, and Sougiannis (2001) and Hirshleifer, Hsu, and Li (2013).

Table 1.6 reports the results. There is no statistically significant relationship between connected holdings and the first two measures of innovation efficiency: patents granted scaled by R&D capital and the total number of citations scaled by R&D capital. But there is a positive and statistically significant relationship between connected holdings and the value created by patents scaled by R&D capital. The results are mixed. To draw a conservative conclusion, connected holdings don’t improve firms’ innovation efficiency. So the observed increases in innovation outcomes are mainly due to more input into innovation.

[INSERT TABLE 1.6 HERE]

1.4.4. Patent Natures

Connected holdings might potentially affect the nature of the patent. I check the impact of connected holdings on the properties of the patents. Trajtenberg, Henderson, and Jaffe (1997) develop the originality and generality measures based on the distribution of citations. A patent that cites diversified technological classes of patents is viewed as more original. On the other hand, a patent that is cited by diversified technologically classes is viewed as more general.

I find that the generality of patents increases when there are more connected holdings. This means that the patents produced by the firms are cited by patents coming from different
technological classes. This evidence is consistent with the higher citation results.

[INSERT TABLE 1.7 HERE]

1.4.5. Real Outcomes

Numerous endogenous growth models imply that firm growth is related to innovation. Since results in the previous sections demonstrate that connected holdings cause an increase in innovation, here I explore whether connected holdings can lead to firm growth. One important caveat is that innovation doesn’t have to be the only channel through which connected holdings affect firm’s growth. Connected holdings might affect the cost of capital of the firm or the firm’s contracting environment. But it is still interesting to investigate whether, besides innovation, connected holdings affect firm’s growth?

I focus on growth of (a) profits, (b) the nominal value of output, and (c) the number of employees. I estimate the following specification:

$$\log Y_{i,t+\tau} - \log Y_{i,t} = \alpha + \beta \text{ConHold}_{i,t} + X'_{i,t} \gamma + \eta_i + \mu_t + U_{i,j,t} + \epsilon_{i,t}, \quad (1.4)$$

where all the variables on the right-hand side are the same as those in the main specification equation (1.2). \(\tau\) is the length of the horizon. Here, I explore \(\tau = 1, 3, 5\).

Table 1.8 reports the results. I find that future firm growth is strongly related to connected holdings. This result is related to Lins, Servaes, and Tamayo (2017). Lins, Servaes, and Tamayo (2017) measure firm social capital by corporate social responsibility (CSR) intensity. They find that during a crisis, high-CSR firms experience higher profitability, growth, and sales per employee relative to low-CSR firms. Here, my finding shows that for a given firm that has more connected holdings, another measure of social capital, they also have higher profit growth, output growth, and employment growth. The results show that connected holdings are, in general, good for the firm.
1.4.6. Robustness

1.4.6.1. Alternative Specification

The patent applications and patent citations for a firm in a certain year are count data, in other words, non-negative integers. To explain count data, linear models have shortcomings: the predicted value of a linear model might be negative, whereas the dependent variable is always positive. For count data, in the literature, people also use a Poisson regression model. My Poisson model specification is as follows:

\[
E(y_{i,t}|x_{i,t}) = \exp(\beta ConHold_{i,t} + X_{i,t}'\gamma + \eta_i + \mu_t + \epsilon_{i,t}).
\]  

(1.5)

Different assumptions about the error term will generate alternative estimators even though equation (1.5) remains the same. A Poisson model assumes the mean equals the variance, but I also consider alternatives, such as a negative binomial regression.

Following Aghion, Van Reenen, and Zingales (2013), I implement the instrumental variable estimator by using the control function approach for the Poisson regression models. For details, please see the appendix.

Table A.6 reports the results. The results are qualitatively similar. Connected holdings lead to more innovation as measured by patents granted and patent citations.

1.4.6.2. Only Focus on the Senior Management

For all the previous analysis, I include all board members. If I restrict my attention to only top management team (defined as CEO, CFO, COO, CTO, and Chairman), do the results hold? Table A.8 reports the results. Comparing columns (1) to (3) with (4) to (6), it is clear that the results are quantitatively similar.
1.4.6.3. Dot-Com Bubble

I explore whether the results are mostly driven by the Internet bubble, since, in those years, the number of patents jumped. Meanwhile, it is the latter part of my sample, so there are more connected holdings (because more people in my data became CEO or a mutual fund manager). So all the results potentially could be driven by those years. Following the literature, I define the Dot-com bubble years as those from 1999 to 2003. I rerun some of the regressions in the two sub-periods (1980 to 1998 and 1999 to 2003), and the results (in untabulated table) are both positive and statistically significant.

1.4.6.4. Management Ownership

Morck, Shleifer, and Vishny (1988) demonstrate that management ownership affects the market valuation of the firm. Innovation, as an important constitute of firm value, might also be influenced by management ownership. Meanwhile, management ownership might also be correlated with connected holdings. To avoid omitted variable bias, I include management ownership as a control variable. I gather management ownership data from the Execucomp Annual Compensation dataset. I construct two measures for management ownership: (1) total ownership by all directors and (2) ownership by CEO. Ownership is defined as the ratio between the shares owned by management and the total shares outstanding.

I include the two additional control variables in the main regression with the total number of patents as the dependent variable. In untabulated tables, I find that the coefficients of connected holdings are still positively statistically significant. And management ownership is also positively statistically significant. These results show that management ownership also matters for innovation.

1.4.6.5. Are Connected Mutual Funds Local Investors?

Coval and Moskowitz (2001) document that local investors have an informational advantage in trading local firms. If all the connected holdings are also local holdings, then it is difficult
to argue that it is the connected mutual fund managers, instead of the local mutual fund managers, who affect the firm’s innovation.

To address this concern, I first compute the localness of the holdings. I obtain the headquarters’ ZIP codes for mutual funds using the CRSP Mutual Fund dataset from WRDS. I obtain the corporate headquarters’ ZIP codes from Compustat. I obtain the mapping between ZIP codes and the corresponding latitude and longitude information from the United States Postal Service (USPS). The location of headquarters is used as opposed to the state of incorporation for the simple reason that firms tend to incorporate in a state with favorable tax, bankruptcy, and takeover laws, rather than for operational reasons, and typically do not have the majority of their operations in their state of incorporation. I use 'Haversine' formula to calculate the great-circle distance between two points—that is, the shortest distance over the earth’s surface.

\[
d = 2R \arctan 2 \left( \sqrt{a}, \sqrt{1-a} \right) \\
a = \sin^2 \left( \frac{\text{lat}_2 - \text{lat}_1}{360\pi} \right) + \cos \left( \frac{\text{lat}_2}{180\pi} \right) \cos \left( \frac{\text{lat}_1}{180\pi} \right) \sin^2 \left( \frac{\text{lon}_2 - \text{lon}_1}{360\pi} \right),
\]

where \(d\) is the distance between two coordinates. \(R\) is the earth’s radius, equals 6,371 km. \text{lat} stands for the latitude, whereas \text{lon} stands for the longitude. The subscripts \(\{1, 2\}\) stand for the places.

[INSERT FIGURE A.4 HERE]

In the data, among all the connected holdings, the local holdings (i.e., a distance of less than 100 km between firms and funds) compose about 9.8%. The distant holdings compose the remaining 90.2%. Next, I repeat the main analysis with the distant connected holdings, and all the results are quantitatively similar.
1.5. Potential Mechanisms

The empirical findings thus far show that more connected holdings lead to more input in R&D and more innovation output, and those effects are stronger in the more innovative industries.

In this section, I discuss several potential explanations for my findings. Since social capital, as measured by connected holdings, might affect multiple dimensions of the firm, all potential stories may co-exist. I focus on three such explanations. First, the literature on corporate short-termism argues that short-term capital market pressure could be detrimental to innovation (Holmström (1989), Stein (1989), and Terry (2016)). I check the impacts of connected holdings on capital market pressures. Second, Manso (2011) predicts that higher job security enables management to take on more innovative projects. Firm exposure to a takeover could be a major job risk to management. I explore the relationship between connected holdings and takeover risks. Third, besides capital market and corporate control market, I investigate the behavior of connected mutual funds in the corporate governance market, especially voting behaviors. If connected funds are more likely to vote in favor of management, this could ease management’s plan in pursuing innovative plans for the firm.

1.5.1. Capital Market Pressures

1.5.1.1. Hypotheses Development

Earning targets matter a lot to the firm. Each fiscal period, public firms must disclose their earnings. Before a disclosure, financial analysts forecast the earnings and the financial press widely reports a consensus forecast for a given firm. This consensus forecast (or earnings target) works as an externally set performance benchmark for the firm. Whether firms beat, maintain, or miss the earnings target matters a lot. Actually, around 90% of surveyed U.S. executives report pressure to meet short-term earnings targets (Graham, Harvey, and Rajgopal (2005)). But why would management care so much about those earnings targets? Previous research (Kasznik and McNichols (2002); Bartov, Givoly, and Hayn (2002)) demon-
strates that missing quarterly earnings benchmarks leads to significantly lower abnormal returns (quarterly and annual). And because of management’s stock-based compensation scheme, low stock returns have a significant impact on management’s compensations.

How does meeting an earnings target matter for innovation? The previous literature shows that firms sacrifice their long-term investments, such as R&D, to meet the short-term earnings targets. It is because the benefits of R&D appear much later, whereas the costs show up in the current quarter through lower earnings. Almost half of surveyed U.S. executives report that they would prefer to reject a positive net present value project over missing their analyst target (Graham, Harvey, and Rajgopal (2005)). The opportunistic cuts to meet short-term targets can be detrimental to innovation.

The role of institutional investors in creating short-termism is mixed. Some institutional investors might exuberate to the managerial myopia behavior because the institutional investors will dump a company’s stock as soon as its earnings are not quite up to par (Stein (1988)). Divestments encourage management to meet targets. However, I hypothesize that the connected mutual fund managers should behave differently. The social capital between the connected mutual fund managers and the firm’s management should encourage the connected mutual fund managers to stay with the firm through adverse situations. So, in this section, I examine the investment behavior of connected mutual funds versus non-connected institutional investors following a missed earnings target.

1.5.1.2. Trading Behaviors

Utilizing the Institutional Brokers Estimates System (IBES) dataset, I document two new stylized facts: (1) when a firm misses its earnings expectations, connected mutual funds stay, whereas non-connected institutional investors divest. (2) Compared with firms with no connected holdings, firms with connected holding see their returns drop less.

My data on analysts’ forecasts come from the IBES database. I focus on the quarterly earnings forecasts because it is widely studied in the finance and accounting literature.
Following the literature, I define an earnings surprise as:

\[
ES_{i,t} = EPS_{i,t} - \text{median}\{EPS^f_{i,k,t}\},
\]

where \(i\) indexes firm, \(t\) indexes time, and \(k\) indexes analyst. \(EPS_{i,t}\) stands for the realized earnings per share. \(EPS^f_{i,k,t}\) is the forecast made by analyst \(k\) for firm \(i\) quarter \(t\)'s earning. \(\text{median}\{EPS^f_{i,k,t}\}\) serves as the consensus forecast for all the analysts.

First, I examine the trading behaviors of connected mutual funds and non-connected funds. To investigate the trading behaviors of different types of investors, I follow a standard event-study methodology. I define trading as the change in the fraction of firm equity from quarter \(t - 1\) to quarter \(t\). For connected holdings, it is

\[
T^c_{i,t} = \text{ConHold}_{i,t} - \text{ConHold}_{i,t-1}.
\]

For non-connected institutional holdings, it is

\[
T^nc_{i,t} = \text{nonConHold}_{i,t} - \text{nonConHold}_{i,t-1}.
\]

For the difference between connected holdings and non-connected institutional Holdings, it is

\[
T^{diff}_{i,t} = (\text{ConHold}_{i,t} - \text{nonConHold}_{i,t}) - (\text{ConHold}_{i,t-1} - \text{nonConHold}_{i,t-1}).
\]

My empirical specification is

\[
y_{i,t} = \text{MISS}_{i,t-1} + \gamma X^t_{i,t-1} + \eta_j + \mu_t + \epsilon_{i,t},
\]
where \( y \) stands for trading by different types of investors and the difference between connected holdings and non-connected holdings. \( MISS_{i,t-1} \) is a dummy variable that equals 1 if firms miss their quarterly earnings target (i.e., \( ES_{i,t-1} < 0 \)). \( X'_{i,t-1} \) is a vector of control variables that includes firm size, book-to-market ratio, and lagged 12 month returns. \( \eta_j \) stands for industry fixed effects. \( \mu_t \) stands for time fixed effects. Standard errors are clustered at the firm level. Table 1.9 presents the results. I find that missing the earnings target triggers a sell off from non-connected institutional investors. The average magnitude is 35 basis points and is statistically significant. For the connected mutual funds, there is no evidence that their trading is affected by these events. And, as a natural result, for the difference between connected holdings and non-connected holdings, the gap becomes wider after firms miss their earning targets. The above evidence shows that after firms miss their earning targets, connected holdings stay and non-connected holdings divest. The results are robust when controlling for firm-level characteristics, as well as time and industry fixed effects.

1.5.1.3. Return Behaviors

The previous section demonstrates that connected holdings are “loyal” to the firm when bad shocks are realized. In this section, I investigate whether connected holdings can actually alleviate the drop in stock price when firms miss their earnings target.

My empirical specification is as follows:

\[
y_{i,t} = \alpha + MISS_{i,t} + AnyCon_{i,t} + MISS_{i,t} \times AnyCon_{i,t} + \epsilon_{i,t},
\]

where \( y \) stands for the next month’s abnormal returns estimated by the following models: CAPM, the Fama-French 3-factor model, the Fama-French-Carhart model, the Fama-French 5-factor model, and the characteristics-adjusted model (DGTW).
The variable of interest is the interaction term $MISS_{i,t} \times AnyCon_{i,t}$, which captures the cross-sectional difference between the firm with connected holdings versus the firm without connected holdings conditioned on firms experiencing a negative earnings surprise.

Table 1.10 presents the result. It is clear that across all kinds of measures of the abnormal return, the coefficient in front of the interaction term is greater than 0. This means that the firms’ return drops less when they are invested with connected holdings. The evidence supports the argument that the presence of connected holdings could actually alleviate the price drop when firms miss the earnings target.

[INSERT TABLE 1.10 HERE]

1.5.2. Management Job Security and Takeover Exposure

In the last section, I show that missed earnings targets lead to lower subsequent stock returns, possibly leading to a reduction in management’s compensation. Another risk associated with missed earnings targets is takeover risk. Stein (1988) argues that “Relatively patient stockholders may not be discouraged by a low earnings report; they may attribute it to a policy of long-term investment by the firm .... Impatient shareholders, on the other hand, may become very distressed by low earnings reports and may try to dump a stock as soon as such a report is issued.... managers will be more fearful of undervaluation and the accompanying possibility of rip-off by a raider”.

Along the chain of reasoning, at its origin, the attitude of the institutional investor is very important. In the last section, I demonstrate that connected mutual funds are more patient with the firm. In this section, I examine whether relationships with connected holdings would lead to lower takeover risk.

I measure the probability of a firm being acquired (i.e., takeover exposure), following Cremers, Nair, and John (2008). I use the expected probability of a firm being acquired instead of actual takeovers because, according to Stein (1988), it is the threat (or likelihood) of a
takeover that affects management’s incentives to invest in innovation.

First, I estimate the firm’s takeover exposure by running a logit model:

\[
Target_{i,t+1} = \frac{1}{1 + e^{-K_{i,t}}} ,
\]

where \( K_{i,t} = \alpha + \beta_1 Q_{i,t} + \beta_2 PPE_{Asset_{i,t}} + \beta_3 \ln Cash_{i,t} + \beta_4 BLOCK_{i,t} + \beta_5 \ln MV_{i,t} \\
+ \beta_6 IndMA_{i,t} + \beta_7 Lev_{i,t} + \beta_8 ROA_{i,t} + \text{Year}_t + \epsilon_{i,t} ,
\]

where \( i \) indexes firm, \( j \) indexes industry, and \( t \) indexes time. The dependent variable \( Target \) is a dummy variable equal to 1 if a firm is a target of an acquisition. All the independent variables are constructed following Cremers, Nair, and John (2008).\(^{16}\) The acquisition data are obtained from the Securities Data Corporations (SDC) database.

After estimating the model using maximum likelihood, I calculate takeover exposure as the predicted probability of the model. Next, I use takeover exposure as the dependent variable in my main specification, like in equation (1.2). I find that connected holdings are statistically negatively associated with takeover exposure. A one-standard-deviation increase in connected holdings leads to a 0.09 (\( = -0.8278 \times 0.0025/0.0222 \)) standard deviation decrease in takeover exposure. This is economically sizable.

According to Stein (1988), if there is less takeover exposure, there is less incentive for management to sacrifice long-term investment for short-term gains. So my findings show that connected holdings could benefit innovation through reducing takeover risk channel.

\[\text{[INSERT TABLE 1.11 HERE]}\]

\(^{16}\)For detailed variable definitions, see the footnote of table 1 in Cremers, Nair, and John (2008).
1.5.3. Connected Funds’ Voting Behaviors

Voting is an important dimension of corporate governance. To continue exploring the idea that connected mutual funds are “supportive” of the board members, I examine connected mutual funds’ voting behaviors. I find that, on average, connected funds are more likely to vote against shareholder-initiated proposals on various governance issues than are non-connected funds.

1.5.3.1. Empirical Test

In 2003, the Securities and Exchange Commission (SEC) introduced a new rule requiring mutual funds to report their votes on all shares held. I gather the data from the Institutional Shareholder Services’ (ISS) Voting Analytics database. This database covers U.S. mutual fund voting records for all institutions filing the SEC N-PX form. It also contains company vote results. I focus on the period from 2005 to 2011.\(^{17}\) Funds have the option of voting “for”, “against”, “abstain”, “withhold”, or “do not vote”. To better assess voter preference, I only focus on the vote with either “for” or “against”.\(^{18}\) Among all the different types of proposals, I focus on four types of shareholder proposals: (1) requires independent board chairman; (2) requires a majority vote for the election of directors; (3) amends articles/bylaws/charter -- calls for special meetings; and (4) restores or provides for cumulative voting. I choose these proposals for two reasons: (1) the board’s attitude is clear; for almost all above mentioned proposals, management recommend the shareholders to vote against it.\(^{19}\) (2) Besides proposals on director election, auditor appointments and executive compensations, these four types of proposals are the most frequently voted proposals.\(^{20}\) I

---

\(^{17}\) I start with 2005 instead of 2003 because only after 2005 (inclusive) can I reliably merge ISS mutual fund voting data with the CRSP mutual fund database. The appendix provides the matching methods. I stop the sample at 2011 because 2010 is the last year in which innovation data are available, and I don’t want to extend the voting analysis in a fashion that might not be relevant for my main (firm innovation) results.

\(^{18}\) “For” and “against” votes compose 95.07% of all mutual fund votes.

\(^{19}\) See the tables in the appendix for management recommendations for these four proposal types.

\(^{20}\) For proposals on director election, please check Iliev and Lowry (2014), Matvos and Ostrovsky (2010), and Cai, Garner, and Walkling (2009). For proposals on executive compensation, see Malenko and Shen (2016) and Butler and Gurun (2012).
merge the voting dataset with the CRSP mutual fund dataset to obtain the fund’s size, expense ratio, turnover ratio, and fund family size. For each firm, I obtain the financial information from CRSP and Compustat. I also obtain institutional holding data from the Thompson Reuters 13F dataset. Lastly, in each year, I label the fund-firm pair connected or not according to the connection definition in Section 1.2.4. The final dataset contains 219,040 mutual fund votes that cover 164 public firms on 250 proposals. There are 8,289 different funds voted.

I explore the impact of connectedness between fund and firm on fund voting behavior. Following the literature, I employ a probit model. Table 1.12 reports the results. The dependent variable is a dummy that equals one if a mutual fund’s vote is consistent with the board recommendation and zero otherwise. The variable of interest is $\text{Connected(broad)}$. It is a dummy that equals one if a mutual fund is connected to firm. For a detailed definition of the connection types, refer to Section 1.2.4. The control variables include ISS recommendation, firm size, book-to-market ratio, the previous year’s total stock return, fund size, fund turnover ratio, fund expense ratio, and fund family size. In some specifications, I also include year and industry fixed effects. To avoid the incidental problem in estimating non-linear model, I use the Fama-French 12-industry classification instead of the 4-digit SIC industry. I cluster standard errors at the fund level.

[INSERT TABLE 1.12 HERE]

Column (1) reports the results for the case without a control. There is a significant positive relationship between connection and support to the management. In column (2), I control for both firm and fund characteristics, as well as industry and year fixed effects. The connection coefficient becomes larger and more statistically significant. In terms of magnitude, the marginal effect of $\text{Connected(broad)}$ is 6.8% (column (2)). This means that if a fund is connected with the firm, it is 6.8% more likely to vote with management on shareholder proposals. In the sample, only 48% of the votes support the management. Holding the marginal effects fixed, if all the funds are connected, the supporting rate increases by 20%.
In columns (3) to (6), I split the sample by specific proposals. The coefficient in front of $\text{Connected(broad)}$ is positive and statistically significant for the first three proposals: requires independent board chairman; requires a majority vote for the election of directors; and calls for special meetings. But it is not significant for the proposal restores or provides for cumulative voting.

In the appendix, I also report the results of the impact of connectedness on firm’s “say-on-pay” proposals in 2011. I find that connected funds, on average, are more likely to vote in favor of management’s proposals on compensation issues.

1.6. Conclusion

In this paper, I provide evidence that firm-level social capital, as measured by connected holdings affect innovation. In particular, I find that more connected holdings lead to more patents, higher patent impacts, and larger firm value as created by the patents. I also find that the above effects are more pronounced for firms in high-tech industries and for connected holdings coming from “truly” active funds. I also find that social capital improves innovation outcomes by encouraging investment in R&D. Connected holdings foster innovation by helping to reduce capital market pressures and to increase management job security.

This work has many exciting future directions. One would be to investigate how institutional investors affect other corporate policies, such as investments and capital structures utilizing the identification strategy developed here. Another would be to investigate the asset pricing implications of social capital. Third, the current setting could be used as a laboratory to examine the impacts of different kinds of networks (neighbor networks, club networks, religion networks, etc.) on corporate real outcomes.
This table reports the first-stage estimation of the instrumental variables analysis. The dependent variable is connected holdings. Two types of connected holdings are studied here. A narrow connection refers to board members and mutual fund managers who are connected by the same university, attended at the same time, received the same type of degree. A broad connection refers to board members and mutual fund managers who are connected by the same university. IV is the total number of connected mutual fund managers to firm $i$ in year $t$ according to the specific connection types. Section 1.2 defines all the variables. Control variables include annual institutional holdings, log sale, log firm age, R&D/Asset, PPE/Asset, leverage, Q, ROA, and CAPX/Asset. The F-stat is cluster adjusted. University FE are defined in Section 1.3.2.2. The estimated model is an ordinary least-squares (OLS) model with fixed effects. Robust standard errors were calculated and are provided in parentheses. Standard errors are clustered at the firm level. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively. I multiply the coefficients for IV by $10^5$ to make them easier to read.
<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>LnPatApp</th>
<th>LnPatCite</th>
<th>LnKPSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ConHold</td>
<td>42.26***</td>
<td>59.47***</td>
<td>8.83***</td>
</tr>
<tr>
<td></td>
<td>(12.83)</td>
<td>(22.60)</td>
<td>(2.10)</td>
</tr>
<tr>
<td>Year FE</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Firm FE</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>University FE</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Control variables</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.84</td>
<td>0.77</td>
<td>0.59</td>
</tr>
<tr>
<td>Observations</td>
<td>110,663</td>
<td>110,663</td>
<td>110,663</td>
</tr>
</tbody>
</table>

This table reports the results of regressions of the innovation outcomes on connected holdings and other control variables. Table A.1 provides variable definitions. Control variables include annual institutional holdings, log sale, log firm age, R&D/Asset, PPE/Asset, leverage, Q, ROA, and CAPX/Asset. University FE are defined in Section 1.3.2.2. The estimated model is two-stage least-squares model with firm and year fixed effects. Robust standard errors were calculated and are provided in parentheses. ***, **, and * indicate that the coefficient is significant at the 1%, 5%, and 10% level, respectively.
Table 1.3: Disaggregating Industry: High-Tech Industry versus the Rest

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>High-tech industry</th>
<th>Non-high-tech industry</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Ln(\text{PatApp})</td>
<td>Ln(\text{PatCite})</td>
</tr>
<tr>
<td>ConHold</td>
<td>101.55***</td>
<td>28.29**</td>
</tr>
<tr>
<td></td>
<td>(37.88)</td>
<td>(11.62)</td>
</tr>
<tr>
<td>Year FE</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Firm FE</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>University FE</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Control variables</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.79</td>
<td>0.86</td>
</tr>
<tr>
<td>Observations</td>
<td>35,534</td>
<td>75,126</td>
</tr>
</tbody>
</table>

This table reports regressions of the innovation outcome on connected holdings for high-tech industries and the remaining industries. In this table, I use the Fama-French 12-industry categorization. The high-tech industry contains: the Fama-French Industry 6 industry classification (Business Equipment -- Computers, Software, and Electronic Equipment) and the Fama-French 10-industry classification (Healthcare, Medical Equipment, and Drugs). Table A.1 provides variable definitions. Control variables include annual institutional holdings, log sale, log firm age, R&D/Asset, PPE/Asset, leverage, Q, ROA, and CAPX/Asset. University FE are defined in Section 1.3.2.2. The estimated model is a two-stage least-squares model with firm and year fixed effects. Robust standard errors were calculated and are provided in parentheses. ***, **, and * indicate that the coefficient is significant at the 1%, 5%, and 10% level, respectively.
Table 1.4: Disaggregating Fund Type: Active versus Closet Indexers

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent variable</td>
<td>LnPatApp</td>
<td>LnPatCite</td>
<td>LnKPSS</td>
</tr>
<tr>
<td>ConHold of active funds</td>
<td>51.85***</td>
<td>77.10***</td>
<td>12.06***</td>
</tr>
<tr>
<td></td>
<td>(16.53)</td>
<td>(28.56)</td>
<td>(3.09)</td>
</tr>
<tr>
<td>ConHold of closet indexers</td>
<td>-9.97</td>
<td>-196.05</td>
<td>-94.43**</td>
</tr>
<tr>
<td></td>
<td>(204.76)</td>
<td>(386.82)</td>
<td>(43.55)</td>
</tr>
<tr>
<td>Year FE</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Firm FE</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>University FE</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Control variables</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.84</td>
<td>0.77</td>
<td>0.55</td>
</tr>
<tr>
<td>Observations</td>
<td>110,663</td>
<td>110,663</td>
<td>110,663</td>
</tr>
</tbody>
</table>

This table reports regressions of the innovation outcomes on different types of connected holdings and other control variables. Active funds are those funds with active share higher than 0.5. Closet indexers are those funds with active shares lower than 0.5. For the definition of active share, please refer to Cremers and Petajisto (2009). Table A.1 provides variable definitions. Control variables include annual institutional holdings, log sale, log firm age, R&D/Asset, PPE/Asset, leverage, Q, ROA, and CAPX/Asset. University FE are defined in 1.3.2.2. The estimated model is a two-stage least-squares model with firm and year fixed effects. Robust standard errors were calculated and are provided in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.
Table 1.5: Connected Holdings and R&D

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ConHold</td>
<td>73.39***</td>
<td>50.46***</td>
<td>63.29***</td>
<td>42.23***</td>
<td>43.53***</td>
<td>7.56***</td>
<td>6.09***</td>
</tr>
<tr>
<td></td>
<td>(22.93)</td>
<td>(16.07)</td>
<td>(16.63)</td>
<td>(13.24)</td>
<td>(16.23)</td>
<td>(1.94)</td>
<td>(1.47)</td>
</tr>
<tr>
<td>Year FE</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Firm FE</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>University FE</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Control variables</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.91</td>
<td>0.93</td>
<td>0.93</td>
<td>0.95</td>
<td>0.94</td>
<td>0.91</td>
<td>0.93</td>
</tr>
<tr>
<td>Observations</td>
<td>52,507</td>
<td>51,749</td>
<td>63,391</td>
<td>62,521</td>
<td>43,470</td>
<td>52,507</td>
<td>51,749</td>
</tr>
</tbody>
</table>

This table reports results of the regressions of the R&D expenditure on connected holdings and other control variables. Table A.1 provides variable definitions. Control variables include annual institutional holdings, log sale, log firm age, R&D/Asset, PPE/Asset, leverage, Q, ROA, and CAPX/Asset. University FE are defined in Section 1.3.2.2. The estimated model is a two-stage least squares model with firm and year fixed effect, and robust standard errors were calculated and are provided in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.
Table 1.6: Connected Holdings and Innovation Efficiency

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Patent/RD</td>
<td>Citation/RD</td>
<td>KPSS/RD</td>
</tr>
<tr>
<td><strong>ConHold</strong></td>
<td>1.96</td>
<td>7.00</td>
<td>20.82***</td>
</tr>
<tr>
<td></td>
<td>(2.43)</td>
<td>(9.41)</td>
<td>(7.31)</td>
</tr>
<tr>
<td>Year FE</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Firm FE</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>University FE</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Control variables</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.41</td>
<td>0.55</td>
<td>0.62</td>
</tr>
<tr>
<td>Observations</td>
<td>110,663</td>
<td>110,663</td>
<td>110,663</td>
</tr>
</tbody>
</table>

This table reports the results of regressions of the innovation efficiency. Definitions of dependent variables can be found in section 1.4.3. Table A.1 provides variable definitions. Control variables include annual institutional holdings, log sale, log firm age, PPE/Asset, leverage, Q, ROA, and CAPX/Asset. The estimated model is a two-stage least-squares model with firm and year fixed effects. Robust standard errors were calculated and are provided in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table 1.7: Connected Holdings and Patent Nature

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Originality</th>
<th>Generality</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ConHold</strong></td>
<td>1.21</td>
<td>4.37*</td>
</tr>
<tr>
<td></td>
<td>(2.29)</td>
<td>(2.42)</td>
</tr>
<tr>
<td>Year FE</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Firm FE</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>University FE</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Control variables</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.41</td>
<td>0.44</td>
</tr>
<tr>
<td>Observations</td>
<td>24,927</td>
<td>23,456</td>
</tr>
</tbody>
</table>

This table reports regressions of the patent Originality and Generality on connected holdings and other control variables. Table A.1 provides variable definitions. Control variables are: annual institutional holdings, log sale, log firm age, R&D/Asset, PPE/Asset, leverage, Q, ROA, CAPX/Asset. University FE are defined in Section 1.3.2.2. The estimated model is a two-stage least squares with firm and year fixed effects. Robust standard errors were calculated and are provided in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.
Table 1.8: Connected Holdings and Firm Growth

<table>
<thead>
<tr>
<th>Panel:</th>
<th>τ = 1</th>
<th>τ = 3</th>
<th>τ = 5</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Profit</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.57</td>
<td>8.70</td>
<td>6.74</td>
</tr>
<tr>
<td></td>
<td>(3.55)</td>
<td>(3.08)</td>
<td>(2.45)</td>
</tr>
<tr>
<td><strong>Panel B: Output</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>5.93</td>
<td>8.17</td>
<td>5.89</td>
</tr>
<tr>
<td></td>
<td>(4.14)</td>
<td>(3.31)</td>
<td>(2.41)</td>
</tr>
<tr>
<td><strong>Panel C: Employment</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>3.23</td>
<td>6.63</td>
<td>5.75</td>
</tr>
<tr>
<td></td>
<td>(2.94)</td>
<td>(2.65)</td>
<td>(2.11)</td>
</tr>
</tbody>
</table>

This table reports regression results for equation 1.4. The dependent variables are firm profit growth, sales growth, and employment growth. τ is the length of the horizon, in unit of years. Table A.1 provides variable definitions. Control variables include annual institutional holdings, log sale, log firm age, R&D/Asset, PPE/Asset, leverage, Q, ROA, and CAPX/Asset. The estimated model is a two-stage least squares model with firm and year fixed effects. t-statistics were calculated and are provided in parentheses.

Table 1.9: Trading Behavior

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MISS</td>
<td>-0.1359</td>
<td>-35.03***</td>
<td>34.89***</td>
<td>13.09***</td>
<td>15.94***</td>
</tr>
<tr>
<td></td>
<td>(0.1490)</td>
<td>(3.531)</td>
<td>(3.537)</td>
<td>(3.665)</td>
<td>(3.504)</td>
</tr>
<tr>
<td>Control variables</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Time FE</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Industry FE</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.0000</td>
<td>0.0008</td>
<td>0.0008</td>
<td>0.0093</td>
<td>0.0098</td>
</tr>
<tr>
<td>Observations</td>
<td>127,879</td>
<td>127,879</td>
<td>127,879</td>
<td>107,858</td>
<td>105,964</td>
</tr>
</tbody>
</table>

The dependent variable in column (1) is the quarterly change of connected holdings; in column (2), it is the quarterly change of non-connected institutional holdings; and, in columns (3) to (5), it is the difference between the quarterly change of connected holdings and the quarterly change of non-connected institutional holdings. MISS is a dummy variable that equals 1 if firms miss their quarterly earnings targets. Control variables include market cap, book-to-market ratio, and the lagged twelve-month total returns. Industry refers to the Fama-French 48-industry classification. Coefficients are reported in basis points. Standard errors clustered at the firm level are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.
<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>(1) $\alpha_{CAPM}$</th>
<th>(2) $\alpha_{F3}$</th>
<th>(3) $\alpha_{FFC}$</th>
<th>(4) $\alpha_{FF5}$</th>
<th>(5) $\alpha_{DGTW}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>MISS</td>
<td>-267.44***</td>
<td>-270.33***</td>
<td>-259.53***</td>
<td>-266.28***</td>
<td>-269.46***</td>
</tr>
<tr>
<td></td>
<td>(8.92)</td>
<td>(8.44)</td>
<td>(8.21)</td>
<td>(8.33)</td>
<td>(9.28)</td>
</tr>
<tr>
<td>AnyCon</td>
<td>14.67</td>
<td>7.66</td>
<td>80.39</td>
<td>0.87</td>
<td>-31.43</td>
</tr>
<tr>
<td></td>
<td>(11.69)</td>
<td>(11.07)</td>
<td>(17.86)</td>
<td>(10.98)</td>
<td>(11.90)</td>
</tr>
<tr>
<td>MISS × AnyCon</td>
<td>90.59***</td>
<td>83.70***</td>
<td>80.39***</td>
<td>79.57***</td>
<td>93.63***</td>
</tr>
<tr>
<td></td>
<td>(19.58)</td>
<td>(18.36)</td>
<td>(17.86)</td>
<td>(17.92)</td>
<td>(19.56)</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.0074</td>
<td>0.0083</td>
<td>0.0079</td>
<td>0.0083</td>
<td>0.0081</td>
</tr>
<tr>
<td>Observations</td>
<td>147,226</td>
<td>147,226</td>
<td>147,226</td>
<td>147,226</td>
<td>147,226</td>
</tr>
</tbody>
</table>

The dependent variable in columns (1) to (5) is the monthly alpha estimated by the following models: CAPM, Fama-French 3-factor model, Fama-French-Carhart model, Fama-French 5-factor model, and the characteristics-adjusted model (DGTW). Dependent variables are in basis points. MISS is a dummy variable that equals 1 if firms miss their quarterly earnings targets. AnyCon is a dummy variable that equals 1 if there is positive connected holdings for firm $i$ in quarter $t$. Standard errors clustered at the firm level are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.
Table 1.11: Connected Holdings and Takeover Exposure

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>Takeover exposure</th>
</tr>
</thead>
<tbody>
<tr>
<td>ConHold</td>
<td>-0.8278***</td>
</tr>
<tr>
<td></td>
<td>(0.1986)</td>
</tr>
<tr>
<td>Year FE</td>
<td>yes</td>
</tr>
<tr>
<td>Firm FE</td>
<td>yes</td>
</tr>
<tr>
<td>University FE</td>
<td>yes</td>
</tr>
<tr>
<td>Control variables</td>
<td>yes</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.8841</td>
</tr>
<tr>
<td>Observations</td>
<td>109,574</td>
</tr>
</tbody>
</table>

The dependent variable is a firm’s takeover exposure. It is computed as the predicted value of the following logit regression:

\[
Target_{i,t+1} = \alpha + \beta_1 Q_{i,t} + \beta_2 \text{PPE/Asset}_{i,t} + \beta_3 \text{LnCash}_{i,t} + \beta_4 \text{BLOCK}_{i,t} + \beta_5 \text{LnMV}_{i,t} + \beta_6 \text{IndMA}_{i,t} + \beta_7 \text{Lev}_{i,t} + \beta_8 \text{ROA}_{i,t} + \text{Ind}_j + \text{Year}_t + \epsilon_{i,t}.
\]

The specification of the regression is the main specification in the paper and can be found in Section 1.3.1. \textit{ConHold} is the annual connected holdings of the firm. Control variables include annual institutional holdings, log sale, log firm age, R&D/Asset, PPE/Asset, leverage, Q, ROA, and CAPX/Asset. University FE are defined in Section 1.3.2.2. Standard errors clustered at the firm level are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.
Table 1.12: Connected Mutual Fund Votes on Governance Issues

<table>
<thead>
<tr>
<th>Dependent variable = 1 if vote with board recommendation</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>All</td>
<td>All</td>
<td>Require independent board chairman</td>
<td>Require a majority vote for the election of directors</td>
<td>Call special meetings</td>
<td>Restore or provide for cumulative voting</td>
<td></td>
</tr>
<tr>
<td>Connected(broad)</td>
<td>0.0940**</td>
<td>0.1728***</td>
<td>0.1763***</td>
<td>0.2805***</td>
<td>0.1974***</td>
<td>0.0374</td>
</tr>
<tr>
<td>Control variables</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Industry FE</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Year FE</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Pseudo R-squared</td>
<td>0.0004</td>
<td>0.1086</td>
<td>0.1531</td>
<td>0.0958</td>
<td>0.0659</td>
<td>0.0535</td>
</tr>
<tr>
<td>Observations</td>
<td>219,040</td>
<td>219,040</td>
<td>63,330</td>
<td>60,097</td>
<td>54,011</td>
<td>41,602</td>
</tr>
<tr>
<td>Marginal effect on Connected(broad)</td>
<td>0.0374</td>
<td>0.0689</td>
<td>0.0635</td>
<td>0.1052</td>
<td>0.0727</td>
<td>0.0147</td>
</tr>
</tbody>
</table>

Each observation represents the vote of a mutual fund on a proposal at a shareholder meeting at a company. The dependent variable equals 1 if mutual fund’s vote follows board’s recommendation. Connected(broad) is a dummy variable that equals 1 if the mutual fund manager attended the same university as at least one of the firm’s board members. Control variables include ISS recommendation, firm size, book-to-market ratio, last year total stock return, fund size, fund turnover ratio, fund expense ratio, and fund family size. Industry refers to the Fama-French 12-industry classification. Standard errors clustered at the firm level are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.
2.1. Introduction

In 2016, active mutual funds in the U.S. managed a total of 11.6 trillion dollars. This industry's annual revenue is on the order of $100 billion, with over one third of this amount representing expenditures on marketing, largely consisting of sales loads and distribution fees (known as 12b-1 fees). Although the relation between mutual fund size, performance, and fees has been actively debated in the academic literature (e.g., Berk and Green (2004), Chen, Hong, Huang, and Kubik (2004), Pástor and Stambaugh (2012), Berk and Van Binsbergen (2015), Pástor, Stambaugh, and Taylor (2015)), the role of marketing and distribution expenditures in steering investors into particular funds is not fully understood. Positive relationship between funds' marketing efforts and flows is well-documented (e.g., Sirri and Tufano (1998a), Barber, Odean, and Zheng (2005), Gallaher, Kaniel, and Starks (2006), Bergstresser, Chalmers, and Tufano (2009), Christoffersen, Evans, and Musto (2013)). Yet marketing expenses contribute substantially to total fund costs, thus reducing returns earned by investors for a given level of fund manager’s skill. Is marketing a purely wasteful rat race, or does it help imperfectly informed investors find attractive investment opportunities more easily (as argued, for example, by Stigler (1961))? Does it enable capital to flow towards more skilled managers or, instead, distort allocation of assets by channeling them towards underperforming funds?

We start with a benchmark model based on Berk and Green (2004), which describes allocation of assets to mutual funds by rational investors in a frictionless market. By estimating the model, we document substantial differences between such an efficient allocation and the observed distribution of fund size. The vast majority of funds are “too big” relative to the model and deliver substantially negative abnormal returns to investors, while the top
decile of funds are actually smaller than is “efficient,” and thus are able to outperform.¹ To explain these differences, we introduce information frictions by generalizing the framework developed by Hortacsu and Syverson (2004) in the context of index funds. We allow funds’ marketing activities, as well as exogenous characteristics, to affect their inclusion in the investors’ information sets. In our setting, both the expense ratios (fees paid by investors) and the marketing/distribution costs (components of these fees related to broker compensation) are endogenous choices of each fund. By estimating our structural model we find that marketing expenses are nearly as important as price (i.e., expense ratio) or performance (i.e., the Bayesian estimate of manager skill based on historical returns) for explaining the observed variation in fund size. Even though marketing in our model is informative, and indeed complementary with fund skill, counterfactual analysis indicates that it reduces welfare due to a positional arm-race that it entails.

We follow Hortacsu and Syverson (2004) and model the impediments to optimal allocation of capital to mutual funds as a search friction, whereby investors randomly sample and evaluate funds until deciding to invest in one of the funds drawn. Heterogeneity in search costs faced by different investors captures the wide variation in financial sophistication (and perhaps even cognitive ability) required to consider and analyze the different investment alternatives. This approach is intuitive, at least when applied to retail investors: the task of choosing among thousands of funds can be daunting even for the most sophisticated individuals, and far more so for those lacking even basic financial literacy.² Mutual fund performance is determined by managerial skill as well as decreasing returns to scale. Investors care about funds’ expected performance and expense ratios, but must learn about latent fund skill from past performance. Hence, our model nests Berk and Green (2004) as a limiting special case where search costs go to zero. Our key innovation is allowing mutual funds to influence the likelihood of being observed by investors through costly marketing (e.g., via broker

¹The bulk of empirical evidence of performance persistence among mutual funds indicates that consistent underperformance is much more prevalent than outperformance - e.g. Carhart (1997).
²Indeed, Bronnenberg, Dubé, Gentzkow, and Shapiro (2015) find that informed retail consumers are less likely to pay the brand premium that is associated with heavily-marketed products.
commissions). Thus, mutual funds choose their expense ratios and marketing expenses, which increase a fund’s probability of being sampled but decrease its profit margin.

We estimate our structural model using data on well-diversified U.S. domestic equity mutual funds. Our estimation results reveal sizable information frictions in the mutual fund market. The average investor implicitly incurs a cost of 39 basis points to “sample” an additional mutual fund. This friction’s magnitude is about 2/3 of the mean annual gross alpha in our sample. The large magnitude of the estimated search cost is a manifestation of the asset misallocation problem that we document. The intuition is simple: high search costs prevent investors from sampling more funds. Less intensive search leads to an inferior allocation, as many investors are forced to “settle” for high-cost, low-skill funds since it is too costly for them to search further (in practice, this could mean investors who lack financial sophistication following a broker’s recommendation without questioning it, consulting other sources, etc.). In comparison, Hortasu and Syverson (2004) find the mean search cost for an average S&P 500 index fund investor is between 11 and 20 basis points. This difference should not be surprising since it is far easier for investors to evaluate index funds (which are essentially identical in terms of the returns they deliver, at least before fees) than actively managed funds. It is also possible that the index fund market is dominated by more sophisticated investors (i.e., those who know to look for them rather than, say, rely on a recommendation of a broker or financial advisor). Our higher estimated search cost indicates that asset misallocation problem is more severe in the mutual fund industry as a whole (including both active funds and passive index funds) than it is within the S&P 500 index fund sector.

Our estimates imply that marketing is relatively useful as a means of increasing fund size. On average, a one basis point increase in marketing expenses leads to a 1% increase in a fund’s size. This effect is heterogeneous across funds. For high-skill funds, it amounts to a 1.15% increase in assets under management, while for low-skill funds it generates only a 0.97% increase. This result is intuitive: since, conditional on being included in an investor’s
information set, a high-skill fund is more likely to be chosen, such funds benefit more from a higher probability of being sampled than low-skill funds. We find that marketing expenses alone can explain 10% of the variation in mutual fund size; this explanatory power is comparable to both fund manager skill and fund price.

We use our model to quantify the importance of marketing expenses and search costs in shaping the equilibrium distribution of fund size as well as its impact on investor welfare via counterfactual experiments. We simulate the impact of preventing funds from doing any marketing by solving for the equilibrium size distribution and funds’ fees choices using the estimated model parameters. We find that if the cap on marketing is set to zero, the mean expense ratio drops from 160 bps in the current equilibrium to 83 bps. Interestingly, funds lower their expense ratios by more than the original amount of marketing costs. The observed average marketing cost is 62 bps, but in the no-marketing equilibrium the average fund price drops by 77 bps. This result indicates that preventing funds from competing on non-price attributes (such as marketing) significantly intensifies price competition. We also find the total share of active funds drops from 74% to 68%. This drop is accompanied by an increase in average fund performance as measured by mean gross alpha. The increase in alpha is due to the effect of decreasing returns to scale on fund performance. In the no-marketing equilibrium, the “index fund” takes up the market share lost by active funds.

Total investor welfare increases by 57% in the counterfactual equilibrium. Three factors contribute to this increase: in the no-marketing equilibrium, (i) active funds are cheaper, (ii) more investors invest in the passive index fund, (iii) active funds’ alpha is on average higher due to the decrease in fund size. The increase in investor welfare is substantially greater than the decline in aggregate mutual fund profits (or even total fee revenue). Thus, in the aggregate marketing reduces welfare. The reason for this inefficiency is that funds engage in a wasteful “arms race” as they compete for the same pool of investors via marketing. For the parameter values that we estimate this effect dominates the positive effect of marketing (steering investors towards more highly skilled funds), even though in general this need not
be true in our model.

In order to further understand the large increase in investor welfare, we examine the cross-section of investor search costs implied by our model. Naturally, high search cost investors search less and pay higher expense ratios than those with low search costs, while the funds they invest in have high marketing fees and lower alphas. Comparing investor welfare in the two equilibria, we show that the bulk of the welfare gain of eliminating marketing is driven by such high search cost investors. The intuition is simple: these are the investors who are “stuck” with the worst funds (unless they happen to be lucky to “find” the index fund or a high-skill active fund). In the no-marketing equilibrium, even the worst funds are much cheaper than in the current equilibrium. This leads to a significant welfare gain for the high-search-cost investors.\(^3\)

In addition, we examine the impact of search costs on equilibrium market outcomes. With the advancement in information technology and development of services enabling more transparent comparison between funds, we would expect the search frictions to decrease over time. We conduct counterfactual experiments where we set the mean search cost to 35 bps and 20 bps respectively. Given a new search cost distribution, funds reoptimize their prices and marketing expenses. We find that as mean search costs decreases from 39 bps to 35 bps, mean marketing expenses drop from 61 bps to 44 bps. But when mean search cost further drops to 20 bps, the equilibrium marketing expenses fall to zero, even though we maintain the regulatory cap at 100 bps. Thus, low search costs render marketing unprofitable. In the model with a high mean search cost, a subset of funds specifically exploit the high search cost investors. Those funds invest aggressively in marketing so as to enter more of the high search cost investors’ choice sets. Since high search cost investors

\(^3\)A potential caveat is that a drastic reduction in marketing expenses could reduce access to financial advice, especially for small investors. If the role of financial advisors is in establishing investors’ trust, as argued by Gennaioli, Shleifer, and Vishny (2015), then investor welfare could be reduced, as would their allocation to the mutual fund sector and, potentially, the equity markets generally. However, empirical estimates by Linnainmaa, Melzer, and Prengler (2018) suggest that net gains to financial advice that trade off certainty equivalent benefit of increased risky asset holdings against the cost of advice are likely to be small.
will not search much, they will invest with those funds. But when mean search cost drops to a sufficiently low level, this strategy is no longer effective. This suggests that our model’s mechanism is consistent with the observed decline in fees charged by active mutual funds along with the growth in passive index funds over the last two decades highlighted by Stambaugh (2014).

While we abstract from the specifics of the interaction between the mutual fund sales force (such as brokers and financial advisors), our model of marketing effort is motivated by a growing literature examining the role of financial advice. Bergstresser, Chalmers, and Tufano (2009) study broker-sold and direct-sold funds and find little tangible benefit of the former to fund investors. Guercio and Reuter (2014) show that the relationship between fund flows and past performance is muted among funds that are sold through brokers, presumably because such funds are targeting investors with higher search costs. Chalmers and Reuter (2012) show that broker recommendations steer retirement savers towards higher-fee funds yielding lower investor returns; Mullainathan, Noeth, and Schoar (2012) provide similar evidence from an audit study of retail financial advisors. Christoffersen, Evans, and Musto (2013) find that the broker incentives impact investor flow to funds, especially for brokers not affiliated with the fund family. Egan, Matvos, and Seru (2016) exhibit the potentially severe conflict of interest between brokers/financial advisors and their retail investor clients, as exemplified by repeat incidence of misconduct in the industry (only about 5 percent of reported misconduct involves mutual funds, however). An alternative to the conflict of interest view is presented by Linnainmaa, Melzer, and Previtero (2016), who show that financial advisors tend to commit common investment mistakes in their own portfolios.

More closely related to our work, Hastings, Hortaçsu, and Syverson (2016) study the role of the sales force on observed market outcomes in the Mexico privatized retirement savings systems. In their model, a fund’s sales force can both increase investors’ awareness of the product and impact their price sensitivity. In our data we cannot distinguish between these
two effects. We thus assume that the observed marketing expenses are purely informative (rather than persuasive). Egan (2017) uses a search-based structural framework similar to ours to study the conflict of interest between brokers and retail investors in the market for structured convertible bonds.

Our paper is related to the literature that aims to understand the observed underperformance of active funds. Pástor and Stambaugh (2012) develop a tractable model of the active management industry. They explain the popularity of active funds despite their poor past performance using two components: decreasing returns to scale and slow learning about the true skill level. In our model of the active management industry, we also include decreasing returns to scale and investor learning about unobserved skill (at the fund level). However, our model largely attributes the popularity of active funds to the information friction that prevents investors from easily finding out about index funds.4

This paper is also related to those studying the role of advertising and media attention in the mutual fund industry. Gallaher, Kaniel, and Starks (2006), Reuter and Zitzewitz (2006), and Kaniel and Parham (2016) study the impact of fund family-level advertising expenditures and the resulting media prominence of the funds on fund flows. In our model, we capture some of these effects parsimoniously by allowing fund family size to impact fund’s probability of being included in investor’s information set.5

The remainder of the paper is organized as follows. Section 2 develops our model. Section 3 describes the data used to estimate the model. Section 4 discusses the estimation methods. Section 5 presents the estimation results. Section 6 conducts the counterfactual analysis. Section 7 concludes the paper.

---

4Huang, Wi, and Yan (2007) argue that differences in mutual fund prominence as well as heterogeneity in the degree of sophistication across investors help explain the observed asymmetry in the response of flows to fund performance. Garleanu and Pedersen (2015) incorporate search costs in their model of active management and market equilibrium, but assume that a passive index is freely available to all investors without the need to search.

5We follow this simple approach to incorporating advertising since the latter constitutes a very small fraction of fund expenditure, compared to the distribution costs that we focus on. Advertising can be potentially quite important for steering consumers into financial products - e.g., Honka, Hortaçsu, and Vitorino (2016) and Gurun, Matvos, and Seru (2016).
2.2. Model

Every period, heterogeneous investors conduct costly search to sample mutual funds to invest their (identical) endowments. Investors care about expected fund performance and expense ratio (i.e., its price). Mutual fund performance is determined by managerial skill as well as the impact of decreasing returns to scale. Mutual funds choose their expense ratios and marketing expenses to maximize profits. Marketing expenditures can increase a fund’s probability of being sampled but decrease its profit margins.

We proceed by first describing how fund’s performance is determined and then the investor’s problem and lastly describe the funds’ behavior.

2.2.1. Fund performance

In a time period $t$, the realized alpha $r_{j,t}$ for an active fund $j \in \{1, 2, ..., N\}$ is determined by three factors: (i) the fund manager’s skill to generate expected returns in excess of those provided by a passive benchmark in that period, denoted by $a_{j,t}$. (ii) the impact of decreasing returns to scale, given by $D(M_t s_{j,t}; \eta)$ where $M_t$ is the total size of the market and $s_{j,t}$ is the market share of the fund $j$, and $M_t s_{j,t}$ denoting fund size, $\eta$ is a parameter measuring the degree of decreasing returns to scale, and (iii) an idiosyncratic shock $\varepsilon_{j,t} \sim N(0, \delta^2)$.

$$r_{j,t} = a_{j,t} - D(M_t s_{j,t}; \eta) + \varepsilon_{j,t}, \quad j = 1, ..., N, \quad (2.1)$$

An important question in the mutual fund literature concerns the relative size of active funds vis-a-vis passive funds (e.g., Pstor and Stambaugh 2012). To be able to address this important extensive margin, we include a single index fund $j = 0$ into our model, and thus abstract from competition between index funds. We assume that the alpha of the index fund is zero, in that it neither has skill nor affected by the decreasing returns to scale. The total market size $M_t$ includes both active funds and the index fund. We treat $M_t$ as an exogenous variable in the model.
Our specification is very similar to Berk and Green (2004) with one exception: the manager’s skill is allowed to vary over time. We assume manager’s skill follows an AR(1) process:

\[ a_{j,t} = (1 - \rho)\mu + \rho a_{j,t-1} + \sqrt{1 - \rho^2} \cdot v_{j,t}, \]  

(2.2)

where \( v_{j,t} \sim \mathcal{N}(0, \kappa^2) \). When a fund is born, its first period skill will be drawn from the stationary distribution \( \mathcal{N}(\mu, \kappa^2) \). Parameter \( \rho \) captures the persistence of the skill level. As with other parameters, its value will be estimated from data. In the limiting case, when \( \rho = 1 \), skill is fixed over time, which is what Berk and Green (2004) assume.

Following Berk and Green, we assume the manager’s skill is not observable to either the investor or fund manager herself: it is treated as a hidden state. Let \( \tilde{a}_{j,t} \) be investor’s belief about the manager’s skill in that period. Since equation (2.2) can be regarded as describing how the hidden state \( a_{j,t} \) evolves over time, and equation (2.1) says that \( r_{j,t} + D(M_t s_{j,t}; \eta) \) is a signal on the hidden state, one can apply Kalman filter to obtain the following recursive formulas for the belief on manager’s skill and the variance of that belief:

\[
\tilde{a}_{j,t} \equiv E( a_{j,t} | r_{j,t-1}, s_{j,t-1}, r_{j,t-2}, s_{j,t-2}, ... ) \\
= \rho \left[ \tilde{a}_{j,t-1} + \frac{\sigma^2_{j,t-1}}{\sigma^2_{j,t-1} + \delta^2} [ r_{j,t-1} + D(M_{t-1} s_{j,t-1}; \eta) - \tilde{a}_{j,t-1} ] \right] + (1 - \rho)\mu, \tag{2.3}
\]

\[
\tilde{\sigma}^2_{j,t} \equiv \text{Var} ( a_{j,t} | r_{j,t-1}, s_{j,t-1}, r_{j,t-2}, s_{j,t-2}, ... ) \\
= \rho^2 \left( 1 - \frac{\sigma^2_{j,t-1}}{\sigma^2_{j,t-1} + \delta^2} \right) \tilde{\sigma}^2_{j,t-1} + (1 - \rho^2)\kappa^2. \tag{2.4}
\]

and \( \tilde{a}_{j,t} = \mu, \tilde{\sigma}^2_{j,t} = \kappa^2 \) for the period \( t \) when \( j \) was born. In the special case of \( \rho = 1 \) these coincide with the expressions derived by Berk and Green (2004) in their Proposition 1. The difference between our updating rule and theirs is that in the Berk and Green (2004) model, all the historical signals receive the same weight in determining the investor’s belief, whereas in our case, when \( \rho \) is smaller than 1, the signals in the more recent periods receive greater weight. This allows us to capture the fact that fund managers and/or their
strategies change over time, and that investors might therefore rationally overweight the recent history.\textsuperscript{6}

2.2.2. Investor search

Each investor allocates a unit of capital to a single mutual fund identified as a result of sequential search (conducted at the beginning of each period \( t \)). Investors are short lived, in the sense that they derive utility from their investment in the fund of their choice, and the capital they invest in the funds dissipates at the end of the period. A new population of investors enters in the subsequent period \( t + 1 \) with new capital endowments that they allocate to the funds, and so on. Let \( p_{j,t} \) be the expense ratio charged by fund \( j \). An investor’s utility derived from investing in fund \( j \) is given by

\[
u_{j,t} = \gamma \tilde{r}_{j,t} - p_{j,t}, \tag{2.5}\]

where

\[
\tilde{r}_{j,t} = \tilde{a}_{j,t} - \eta \log(M_t s_{j,t}).
\]

Recall that \( \tilde{a}_{j,t} \) is the investors’ belief on the manager’s skill for fund \( j \) for this period \( t \) and \( \tilde{r}_{j,t} \) is the fund \( j \)'s expected alpha in period \( t \) implied by these updated beliefs as well as the size of the fund, given the decreasing returns to scale function parameterized as \( D(M_t s_{j,t}; \eta) = \eta \log(M_t s_{j,t}) \). The coefficient in front of the expense ratio is normalized to 1. If \( \gamma = 1 \) then investors simply care about the expected outperformance net of fees (net alpha), as assumed by Berk and Green (2004). We allow the more general formulation to account for the potential difference in salience of fees vs. performance as well as the investors’ imperfect ability to estimate manager skill. The utility derived from investing in the index fund is given by \( u_{0,t} = -p_{0,t} \), where \( p_{0,t} \) is the expense ratio charged by the index fund in period \( t \); the alpha of the index fund is set to be zero.

\textsuperscript{6}There is evidence that investors “chase” recent performance, potentially more actively than would be justified from a purely Bayesian perspective. Our framework could be used in quantitatively assessing the degree to which this behavior is driven by irrational over-extrapolation - e.g. Bailey, Kumar, and Ng (2011), Greenwood and Shleifer (2014).
Fix a time period $t$. Investor $i$ pays search cost $c_i$ to sample one fund from the distribution of funds (this distribution is known to all investors, while specific fund identities are not). The search costs are different across the population of investors and follow a continuous distribution $G$ (which we parameterize as exponential, with its mean given by $\lambda$). As in Hortasu and Syverson (2004), we endow investors with one free search, so that every investor will invest in a fund (even if his search cost is very high). Let $\Psi_t(u)$ be the probability of sampling a fund that delivers the investor an indirect utility smaller or equal to $u$. Standard Bellman equation arguments imply that it is optimal for the investor to follow a cutoff strategy (see Appendix for details). Let $u^*$ be the highest indirect utility among the funds sampled thus far. The investor continues searching iff $u^* \leq \bar{u}(c_i)$, where the threshold $\bar{u}$ is defined by

$$c_i = \int_{\bar{u}}^{+\infty} (u' - \bar{u}) d\Psi_t(u').$$

Since we have a finite number of funds, the above expression becomes

$$c_i = \sum_{j=0}^{N} \psi_{j,t}(u_j - \bar{u}) \cdot 1\{u_j > \bar{u}\},$$

where $\psi_{j,t}$ is the probability of sampling fund $j \in \{0, 1, ..., N\}$. Intuitively, the left hand side is the cost for an additional search, and the right hand side is the expected gain. Note that the right hand side is strictly decreasing in $\bar{u}$. So $\bar{u}(c_i)$ is strictly decreasing in $c_i$. Intuitively, the bigger $c_i$ is, the smaller the cut-off $\bar{u}(c_i)$ becomes, and the less persistent the investor is in searching. Following Hortasu and Syverson (2004), we can solve for the market share of each fund, $s_{j,t}$, explicitly as a function of the utilities $\{u_{j,t}\}_{j=0}^{N}$, sampling probabilities $\{\psi_{j,t}\}_{j=1}^{N}$ and the distribution of search costs $G(c_i)$ (see detailed derivations in the Appendix).
### 2.2.3. Marketing and equilibrium market shares

Fund sampling probabilities depend on fund characteristics and, crucially, on funds’ marketing efforts. Let $b_{j,t}$ denote marketing expenses of fund $j$, $x_{j,t}$ denote a vector collecting the (observable) exogenous characteristics of the fund, and $\xi_{j,t}$ represent the unobservable shock that affects the sampling probability of this fund. Vector $x_{j,t}$ includes year dummies, fund age, and the number of funds in the same family. Then the probability that an investor randomly draws fund $j$ in year $t$ is specified as

$$
\psi_{j,t} = \frac{e^{\theta b_{j,t} + \beta' x_{j,t} + \xi_{j,t}}}{1 + \sum_{k=1}^{N} e^{\theta b_{k,t} + \beta' x_{k,t} + \xi_{k,t}}},
$$

(2.6)

$$
\psi_{0,t} = 1 - \sum_{k=1}^{N} \psi_{k,t}.
$$

(2.7)

Thus, $\theta$ is a key parameter that characterizes the effectiveness of marketing expenditure as a means of attracting investors. As long as $\theta$ is positive, an increase in $b_{j,t}$ increases the probability that the fund is sampled by investors, all else equal. We assume that the index fund does not engage in any marketing activities; thus, increasing marketing by all of the active funds automatically reduces its sampling probability.

Importantly, marketing expenditures are only incurred by the fund when it attracts investment. This is intuitive if we think of marketing as a commission paid to a broker or advisor - while a promise of a kickback might increase the probability the fund is recommended, if the client chooses not to invest in the fund, the commission is not paid. While we do not model this intermediated relationship between investors and funds in detail, our specification of the sampling probability function is meant to capture it in a somewhat reduced form.\(^7\) In addition, it captures other aspects of marketing that are not directly tied to the marketing expenses that reported a part of the fund expense ratio, such as “brand,” which might be captured by both observed and unobserved characteristics of the fund.

\(^7\)The role of commissions/kickbacks in investment management is analyzed theoretically by Inderst and Ottaviani (2012) and Stoughton, Wu, and Zechner (2011).
We use $p_t$ to denote the vector that collects $p_{jt}$ for $j = 1, ..., N$; similar notation applies to other fund-specific variables in the model. With the specifications in (2.5), (2.6), and (2.7), the search model in Section 2.2.2 implies a mapping from $p_t$, $b_t$, $\tilde{r}_t$, $x_t$, $\xi_t$, and $p_{0,t}$ to a set of market shares. Let us write this mapping as

$$s_{jt} = F_{jt}(p_t, b_t, \tilde{r}_t, x_t, \xi_t, p_{0,t}; \Theta), \quad j = 1, ..., N,$$  

(2.8)

where $\Theta$ collects the relevant parameters, which in this case include $\gamma$, $\beta$, $\theta$, and the parameter $\lambda$ for $G$. The share for the index fund is given by $s_{0,t} = 1 - \sum_{j=1}^{N} s_{jt}$. We use vector $s_t$ to collect $s_{jt}$ for $j = 1, ..., N$.

Decreasing returns to scale imply that $\tilde{r}_t$ depends on the funds’ market shares:

$$s_t = F_t[p_t, b_t, \tilde{a}_t - \eta \log (M_t s_t), x_t, \xi_t, p_{0,t}; \Theta].$$  

(2.9)

As in Berk and Green (2004), investors understand that their returns depend on the size of the fund they invest in, and therefore the equilibrium vector of fund market shares $s_t$ is a fixed point of the above relation. We can write the fixed point as a function of the other inputs on the right hand side of (2.9),

$$s_{jt} = H_{jt}(p_t, b_t, \tilde{a}_t, x_t, \xi_t, p_{0,t}; \Theta), \quad j = 1, ..., N,$$  

(2.10)

with $\Theta$ now also including parameter $\eta$. In the appendix, we show that this fixed point is unique. Unlike $F_{jt}$, we do not have a closed-form expression for $H_{jt}$ and so it requires fixed point iteration to compute.

Profits for an active fund $j$ in period $t$ are given by

$$\pi_{jt} := M_t \cdot H_{jt}(p_t, b_t, \tilde{a}_t, x_t, \xi_t, p_{0,t}; \Theta) \cdot (p_{jt} - b_{jt}).$$  

(2.11)

We assume a Nash equilibrium, where each fund chooses $p_{jt}$ and $b_{jt}$ to maximize $\pi_{jt}$, given
other funds’ choices \( p_{-j,t} \) and \( b_{-j,t} \) (as well as its own and other funds’ estimated skills and exogenous characteristics).\footnote{Our notion of profits most closely approximates management fees paid to the fund’s investment advisor. We can think of this either as profits accruing to the fund family or as compensation paid to the fund manager, although in reality the latter is a much more complicated object, e.g., see Ibert, Kaniel, Van Nieuwerburgh, and Vestman (2017).} Because the SEC currently imposes a one-percent upper bound on the 12b-1 fees, we restrict \( b_{j,t} \leq \bar{b} \equiv 0.01 \) in equilibrium.

In sum, in our model mutual funds choose their fees and marketing efforts to maximize profits each period while taking into the equilibrium distribution of fund size (and therefore expected outperformance of each fund).

2.3. Data

The data come from CRSP and Morningstar. Our sample contains 2,285 well-diversified actively managed domestic equity mutual funds from the United States between 1964 and 2015. Our dataset has 27,621 fund/year observations. In the data appendix, we provide the details about how we construct our sample. We closely follow data-cleaning procedures in Berk and van Binsbergen (2015) and Pstor, Stambaugh and Taylor (2015).

To compute the annual realized alpha \( r_{j,t} \), we start with monthly return data. We first augment each fund’s monthly net return with the fund’s monthly expense ratio to get the monthly gross return \( r_{j,t}^{\text{Gross}} \). Then we regress the excess gross return (over the 1-month U.S. T-bill rate) on the risk factors throughout the life of the fund to get the betas for each fund. We multiply betas with factor returns to get the benchmark returns for each fund at each point in time. We subtract the benchmark return from the excess gross return to get the monthly gross alpha. Last, we aggregate the monthly gross alpha to the annual realized alpha \( r_{j,t} \). We use 4 different benchmark models: CAPM, Fama-French three-factor model, Fama-French and Carhart four-factor model and Fama-French five-factor model. For our main results, we use the Fama-French five-factor model as the benchmark, but our results are robust to other risk adjustments. In our sample, the average annual realized alpha for Fama-French five-factor model is 54 bps. This result is very close to Pstor, Stambaugh and
Taylor (2015)'s estimates, where they find the monthly alpha is 5 bps, which translates to 60 bps of annual alpha.

Since our focus is on the efficient allocation of assets across active funds, we choose to minimize the details related to modeling index funds. We aggregate all index funds from Vanguard to build a single index fund. We choose Vanguard because, as argued by Berk and van Binsbergen (2015), Vanguard index funds have been the most accessible index funds for retail investors historically. Specifically, we compute its assets under management (hereafter AUM) by summing AUM across all funds; we compute the combined fund's expense ratio by asset-weighting across all included index funds. We count the combined index fund's age from the inception year of Vanguard, which is 1975.

We define the total mutual fund market $M_t$ as the sum of AUMs of all the active funds and the combined index fund in year $t$. We define market share $s_{j,t}$ as the ratio between fund $j$'s AUM and the total fund market. $M_t s_{j,t}$ gives the fund $j$'s AUM in millions of dollars in year $t$. We exclude fund/year observations with fund's AUM below $15$ million in 2015 dollars. A $15$ million minimum is also used by Elton, Gruber, and Blake (2001), Chen, Hong, Huang, and Kubik (2004), Yan (2008), and Pástor, Stambaugh, and Taylor (2015). In our dataset, there is a huge skewness in fund’s AUM. From the summary statistics, we can see the mean of fund’s AUM is much larger than the median. The funds at the 99 percentile is over 1,100 times larger than the funds at the 1 percentile. This skewness could potentially affect our estimates. Following Chen et al (2004) we use the logarithm of a fund’s AUM as our measure of fund size.

In taking our model to the data, we use the reported distribution costs (sales loads and 12b-1 fees, which are typically used to compensate brokers for directing client investment to funds) as our combined proxies for marketing costs (rather than using, for example, advertising expenditures). The reason is that in the U.S., many investors purchase mutual funds through intermediaries such as brokers or financial advisors. Among all the expenses

---

9For a detailed study of search frictions within the index fund market, see Hortasu and Syverson (2004).
that mutual fund companies categorized as marketing, advertising expenses constitute only a tiny portion (according to the ICI). The bulk of the marketing costs is compensation paid to brokers and financial advisors, albeit we do not observe this compensation directly.

In the mutual fund industry, a single mutual fund may provide several share classes to investors that differ in their fees structures (typically, the difference is in the combination of front loads and 12b-1 fees). Following much of the literature (with some exceptions, e.g., Bergstresser, Chalmers, and Tufano 2009), we conduct our analysis at the fund level instead of the share class level. To be able to do so, we need to aggregate the share class level expense ratios, 12b-1 fees and front loads up to the fund level. We define the marketing expense $b_{j,t}$ as the “effective” 12b-1 fees that includes amortized loads (see appendix for details).

2.4. Estimation

Our estimation proceeds in two steps. We first estimate the set of parameters governing mutual fund investment performance: $\mu$, $\kappa$, $\delta$, $\rho$, and $\eta$, using the observed panel of fund returns and market shares: $\{r_{j,t}, s_{j,t}|j = 1, ..., N, t = 1, ..., T\}$ using maximum likelihood estimation (MLE). This also gives us the posterior beliefs on the funds’ skills in every period. Then we estimate the other parameters (which are related to the search model) using generalized method of moments (GMM) by relating the observed $s_{j,t}$ to the fund characteristics, as well as making inferences from the equilibrium restrictions on the fee-setting behavior of funds, taking the Bayesian posterior beliefs about funds’ skills as given.

2.4.1. Fund performance

From expression (2.1), we can write down the probability of observing $r_{j,t}$ conditional on observed market shares and realized outperformances up to $t$:

$$\Pr \left( r_{j,t} | s_{j,t}, r_{j,t-1}, s_{j,t-1}, r_{j,t-2}, s_{j,t-2}, ... \right) \sim N \left[ \tilde{a}_{j,t} - \eta \log(M_t s_{j,t}), \tilde{\sigma}_{j,t}^2 + \delta^2 \right].$$
In writing down the above conditional likelihood, note that the current market share \( s_{j,t} \) does not provide further information about the skill \( a_{j,t} \) beyond \( \{r_{j,t-1}, s_{j,t-1}, r_{j,t-2}, s_{j,t-2}, \ldots\} \), because it is a function of \( \tilde{a}_{j,t} \) but not \( a_{j,t} \) directly. Neither does \( s_{j,t} \) provide any information on \( \varepsilon_{j,t} \) for the same reason.

We can use the above conditional probability to construct a partial log likelihood function (see Wooldridge (2010), § 13.8):

\[
\sum_{j=1}^{N} \sum_{t} \log \Pr \left( r_{j,t} \mid s_{j,t}, r_{j,t-1}, s_{j,t-1}, r_{j,t-2}, s_{j,t-2}, \ldots \right).
\]

The first summation is across all funds. The second summation is across all periods in which fund \( j \) exists. We maximize this likelihood with respect to \( \mu, \kappa, \delta, \rho, \) and \( \eta \) to obtain the estimates. Our MLE estimation does not rely on the assumptions and structure of our search model - only on the specification of the exogenous skill process. Therefore our estimates of skill and decreasing return to scale parameters are valid even if fund market shares are determined by some other model, for example Berk and Green (2004), which we use as a benchmark.

### 2.4.2. Search model

The parameters in the search model are estimated using (i) a set of moment conditions constructed with \( \xi_{j,t} \) and (ii) the optimality of the funds’ behaviors. For (i), we first need to back out the \( \xi_{j,t} \)'s from the data given any set of parameter values. This amounts to finding the \( \xi_t \) that equates the model-predicted market shares \( H_t(p_t, b_t, \tilde{a}_t, x_t, \xi_t, p_{0,t}; \Theta) \) to the observed shares \( s_t \) for each period \( t \). Since the fixed point of \( H_t \) is observed as \( s_t \) in the data we can achieve this by solving \( F_t[p_t, b_t, \tilde{a}_t, \eta \log (M_t s_t), x_t, \xi_t, p_{0,t}; \Theta] = s_t \) (which is described in equation (2.9)) for \( \xi_t \) (given a set of parameter values and observed fund choices).\(^{10}\)

\(^{10}\)The solution to this equation can be found by iteration similar to the contraction mapping approach in Berry, Levinsohn, and Pakes (1995).
The definition of $\xi_t$ gives us our first set of moment conditions: $E(\xi_{j,t}|x_{j,t}, \tilde{a}_{j,t}) = 0$. This condition states that $\xi_{j,t}$ is mean independent of the $x_{j,t}$, the exogenous variables that affect fund’s sampling probability in addition to marketing expenses, and $\tilde{a}_{j,t}$, the posterior belief about the fund skill at the beginning of period $t$. Let $j \in t$ denote any active fund $j$ that is alive in period $t$. The sample version of the moment conditions is

$$
\sum_{t=1}^{T} \sum_{j \in t} \xi_{j,t} \begin{pmatrix} x_{j,t} \\ \tilde{a}_{j,t} \end{pmatrix} = 0.
$$

Following Hortasu and Syverson (2004) and Chen et al. (2004) we include in $x_{j,t}$ both log age and the number of funds in the same fund family to capture the fund level social learning effects, as well as advertising that is conducted at the family level. Importantly, we do not include lagged fund size into $x_{j,t}$. In the data, fund size is persistent over time, so including lagged fund size creates an over-fitting problem, where $s_{j,t}$ is almost mechanically explained by $s_{j,t-1}$. From the point of view of the moment conditions, such a problem arises due to the fact that lagged size depends on $\xi_{j,t-1}$, which is also likely persistent, and so lagged size $s_{j,t-1}$ is likely correlated with $\xi_{j,t}$.

In contrast to this first set of moment conditions, we do not require $E(\xi_{j,t}|p_{j,t}, b_{j,t}) = 0$ because $p_{j,t}$ and $b_{j,t}$ are endogenous outcomes of the model and thus depend on $\xi_{j,t}$. One typical approach that the literature explores to deal with such endogeneity is using instruments for firms’ pricing or marketing choices. Another common approach, which we follow here, is to rely on the optimality of the observed firm choices. Intuitively, the levels of fees and marketing expenses that are optimal for different funds will depend on the elasticities of demand. Therefore, as long as the observed choices are optimal, they help to identify the demand function.

The first order condition for the price for fund $j$ at period $t$ is

$$
s_{j,t} + \frac{\partial H_{j,t}}{\partial p_{j,t}} \cdot (p_{j,t} - b_{j,t}) = 0.
$$

67
In order to exactly align the behaviors predicted by a model with the observed behaviors of each individual fund in the data, one must either introduce unobserved heterogeneity in costs or allow for the first-order conditions to be satisfied with error.\textsuperscript{11} In our estimation, we implicitly allow decision errors as each fund chooses its price and marketing expense. Specifically, we allow the first-order condition above to be satisfied up to an error:

\begin{equation}
    s_{j,t} + \left( \frac{\partial H_{j,t}}{\partial p_{j,t}} \cdot e^{\zeta_{j,t}} \right) (p_{j,t} - b_{j,t}) = 0, \tag{2.12}
\end{equation}

where $\zeta_{j,t}$ represents the fund’s “error” in setting its price (e.g., due to a mis-assessment of the slope of the demand curve). We will assume that $\zeta_{j,t}$ has a mean of zero across all periods and funds. In other words, while discrepancies are allowed at the individual fund level, we still ask the average behavior to be consistent with the model.

The first order condition for the marketing expenses is similar but sightly more involved because of the corner restrictions $0 \leq b_{j,t} \leq \bar{b}$. We let

\begin{equation}
    -s_{j,t} + \left( \frac{\partial H_{j,t}}{\partial b_{j,t}} \cdot e^{\omega_{j,t}} \right) (p_{j,t} - b_{j,t}) \begin{cases} 
        \leq 0, & \text{if } b_{j,t} = 0; \\
        \geq 0, & \text{if } b_{j,t} = \bar{b}; \\
        = 0, & \text{otherwise.}
    \end{cases} \tag{2.13}
\end{equation}

Here we again allow a mean zero error $\omega_{j,t}$. One interpretation of these decision errors is inertia: if it is costly for funds to change the fees that they charge, including the component that covers marketing costs, these will be sticky over time, typically deviating from the level that is optimal at a particular point in time (we abstract from modeling the dynamic fee-setting behavior here). Another source of errors would come from fund-family-related constraints (e.g., some families have financial advisory arms and might choose to cross-subsidize those by channeling marketing fees to the advisors they employ, even if it is

\textsuperscript{11}See Baye and Morgan (2004), which shows that allowing only a small amount of bounded rationality in players’ optimization behaviors can be of great use in reconciling the Nash hypothesis with the commonly observed price patterns in the data.
suboptimal from the standpoint of maximizing profits on some of the funds they manage, while other families might eschew marketing altogether even if some of their funds might benefit from it).\textsuperscript{12}

Thus, we assume that fund choices of prices and marketing expenses are optimal on average, so that the first order conditions are satisfied up to a fund-period-specific errors. Given (2.12) and (2.13), this amounts to $E(\zeta_{j,t}) = 0$ and $E(\omega_{j,t}) = 0$. Notice that these moments do not impose any distributional or correlational assumptions on $\zeta_{j,t}$ or $\omega_{j,t}$. In sample versions,

\begin{align*}
\sum_{t=1}^{T} \sum_{j \in t} \zeta_{j,t} &= 0, \quad \text{(2.14)}
\sum_{t=1}^{T} \sum_{j \in t} \omega_{j,t} &= 0. \quad \text{(2.15)}
\end{align*}

The first error, $\zeta_{j,t}$, can be directly backed out from the first order condition given any set of parameter values:

$$\zeta_{j,t} = -\log \left( \frac{-\partial H_{j,t}/\partial p_{j,t}}{s_{j,t}} \right) - \log (p_{j,t} - b_{j,t}).$$

The second error, $\omega_{j,t}$, can be computed exactly for $j$ with $0 < b_{j,t} < \bar{b}$, but unfortunately not for the boundary cases:

$$\omega_{j,t} \begin{cases} 
\leq \bar{\omega}_{j,t}, & \text{if } b_{j,t} = 0; \\
\geq \bar{\omega}_{j,t}, & \text{if } b_{j,t} = \bar{b}; \\
= \bar{\omega}_{j,t}, & \text{otherwise},
\end{cases}$$

\textsuperscript{12}There is a connection between the decision errors that we introduce here and the notion of $\epsilon$-equilibrium in game theory, first introduced by Radner (1980). A set of choices constitutes an $\epsilon$-equilibrium if the difference between what a player achieves and what he could optimally achieve is less than $\epsilon$. In other words, it only requires each player to behave near-optimally, which turns out to be the same as what we ask in (2.12) and (2.13). Specifically, there is a mapping from $\zeta_{j,t}$ and $\omega_{j,t}$ to the loss that firm $j$ incurs relative to its optimal payoff. When both errors are zero, such loss is zero. More importantly, it can be shown that this mapping is insensitive, in the sense that fairly large errors only lead to a relatively small loss reduction in profits.
where
\[ \bar{\omega}_{j,t} \equiv -\log \left( \frac{\partial H_{j,t}}{\partial b_{j,t}} \frac{s_{j,t}}{\partial b_{j,t}} \right) \log (p_{j,t} - b_{j,t}). \]

In principle, we cannot simply use the average of \( \bar{\omega}_{j,t} \) as an estimate of \( E(\omega_{j,t}) \). A conventional way to deal with this kind of truncation problem is to make an additional distributional assumption and apply an MLE estimator. However, a key issue here is that the truncation interval is not fixed but varies across funds endogenously, and thus it may be correlated with \( \omega_{j,t} \).

We take a simpler approach of comparing the estimates based on several subsample versions of (2.15):

(i) \[ \sum_{0 < b_{j,t} < \bar{b}} \omega_{j,t} = 0; \] (ii) \[ \sum_{b_{j,t} = 0} \bar{\omega}_{j,t} = 0; \] (iii) \[ \sum_{b_{j,t} = \bar{b}} \bar{\omega}_{j,t} = 0; \] (iv) \[ \sum_{all \ j, t} \bar{\omega}_{j,t} = 0. \]

The first version (i) assumes that on average, the funds that choose an interior level of marketing expenditure are right about the effect of marketing on market share. These are the funds for which we can exactly calculate the \( \omega_{j,t} \). We acknowledge that these funds are a selected sub-sample of all funds; their average does not necessarily reflect the average across all funds. However, these are the funds that choose the less extreme marketing expenses. In addition, they make up a substantial portion (about 30%) of the funds in the data, so it is reasonable to believe that their average assessment is not far from the population average.

The second version (ii) uses the truncated values (lower bounds) of the \( \omega_{j,t} \) of the funds that choose zero broker marketing expenses. The third version (iii) uses the truncated values (upper bounds) of the \( \omega_{j,t} \) of the funds that choose the highest possible marketing expenses, \( \bar{b} \), which has been 1 percent imposed by the SEC. The last version (iv) uses all the values for \( \omega_{j,t} \). We use these three latter cases as robust checks. If the estimates based on these four different assumptions are similar, then we can be confident that estimates based on the full sample moment (2.15) are not too severely impacted by the truncation.\(^{13}\)

\(^{13}\) We analyze how sensitive our parameter estimates to the violations of these two key moment conditions using the method proposed by Andrews, Gentzkow, and Shapiro (2017) in the Appendix. We show that
Our GMM estimation is just identified, since there are five unknown parameters (not counting the year fixed effects) and five moment conditions. The parameters are the average search cost $\lambda$, the utility weight of outperformance $\gamma$, the sensitivity of sampling probability to marketing $\theta$, and a two-dimensional vector of sensitivities $\beta$ (for number of funds in the fund family and fund age). There are three moment conditions for the sampling probability residual $\xi$ and two more moment conditions based on the first order necessary conditions for the optimality of funds’ pricing and marketing behavior in equations (2.14) and (2.15), respectively. We conduct this second-stage estimation in one step using the identity weighting matrix. Standard errors are estimated via parametric bootstrap (described in the Appendix), which allows for arbitrary correlation between error terms, in particular $\zeta_{j,t}$ and $\omega_{j,t}$.

2.5. Results

2.5.1. Fund performance

Table 2.1 reports estimates of the fund performance-related parameters using our full sample. The magnitude of decreasing returns to scale parameter $\eta$ is 0.0048, and it is statistically significant. Since one standard deviation of log fund size is 1.628, a one standard deviation increase in log fund size is associated with approximately 78 basis points decrease in mean annual alpha. This result is close to Chen et al. (2004). This magnitude is economically significant, in particular as compared to the mean gross alpha of 54 basis points. For robustness, we also estimate the model using linear rather than logarithmic specification similar to Pstor, Stambaugh and Taylor (2015) and obtain estimates broadly consistent

reasonable deviations from the assumption that these moment conditions hold with equality, even at the boundaries, implies negligible changes in the estimated parameter values.

In our dataset, the first period with non-missing data is the year 1964, so our full sample estimates use the data from 1964 to 2015. It is important to use all the available information to estimate the learning model that is the core of Berk and Green (2004). In the model, when a fund is born, it draws an initial skill level from the prior skill distribution. Then investors use the entire history of subsequent realized performance to update their beliefs about each fund’s skill level. If we were to start the sample at a later date, for example, year 1995, we would lose the performance information for a lot of funds that were in operation well before 1995. One way to circumvent the above truncation problem is to pick a starting year and keep only the funds which are founded after this year. But this approach would bias the estimates toward newer funds.
Existence of stock-picking skill among mutual fund managers is one of the oldest queries in empirical finance. Early literature used persistence of fund-level performance as an indicator of skill in active fund management, an approach that is called into question by Berk and Green (2004), who argue that the lack of performance persistence does not imply absence of fund manager skill, as long as capital flows to outperforming funds and if fund size erodes performance. Here we take a different approach by estimating a version of the Berk and Green model (2004) directly. We find that the mean of the prior distribution of managerial skill is 3.05% (per annum). This number is positive and statistically significant, which means that an average active mutual fund manager is skilled (we plot the prior distribution of fund manager skill implied by the estimated parameters $\mu$ and $\kappa$ in Figure A.8 in the Appendix). Over 71% of the funds have fundamental skill levels that are higher than the mean expense ratio, at least when applied to the first dollar of assets under management (i.e., before any of the effects of decreasing returns to scale).

Another parameter of interest is $\rho$, the persistence of fund manager’s skill. Our empirically estimated persistence is 0.94, which means past beliefs are quite useful in predicting future performance. Our skill persistence result is consistent with Berk and van Binsbergen (2015) who find that the cross sectional differences in value added persist for as long as 10 years. One of the reasons that skill persistence is not perfect as assumed in the Berk and Green model is managerial turnover. If we believe the skill of a mutual fund is partially due to the mutual fund manager, then a change of management team might affect the skill level of the fund. Fidelity Magellan fund manager Peter Lynch is a case in point: during his tenure from 1977 to 1990, according to our measure of performance, Magellan fund achieved 14

---

15There is some disagreement in the literature regarding the role of fund size in eroding performance due to decreasing returns to scale, since estimates of the latter can potentially suffer from omitted variable bias, since both fund size and observed performance are correlated with underlying fund skill, which is unobservable - e.g., see discussion in Pstor, Stambaugh and Taylor (2015) and Reuter and Zitzewitz (2015). We verify that our estimation approach can recover the true value of the parameter that controls the strength of decreasing returns without noticeable bias via Monte Carlo simulation. The apparent absence of a meaningful bias is in part due to the fact that in the data, as we show below, fund size is only weakly related to fund manager skill.
consecutive years of positive alpha. After Peter Lynch’s departure, Magellan’s performance becomes less impressive, reverting towards the mean.

2.5.2. Asset misallocation

Equipped with estimated parameters of fund skill distribution, we compute the investor beliefs about each fund’s skill level at each point in time. Then we can derive implied fund size according to a benchmark frictionless model following Berk and Green (2004) (henceforth BG). By comparing BG-implied fund size with the data, we can assess the degree of asset misallocation in the mutual fund industry.

First, we compute the investor beliefs about each fund’s skill level $\tilde{a}_{j,t}$ using the recursive expression derived in section 2.2.1. As an example, consider a fund $j$ that was born in period $t = 1$. At the fund’s birth, we assign the fund an expected skill level of $\mu$, then we use realized return $r_{j,1}$ and fund size $M_1s_{j,1}$ to get the updated belief, $\tilde{a}_{j,1}$. By iterating forward, we can generate the whole series of fund’s expected skill levels. Next, we compute the BG-implied fund size. Berk and Green’s model predicts that fund’s size (i.e., total assets under management), which we denote by $s_{j,t}^{BG}$, should be such that the decreasing returns to scale exactly offsets the investor belief less fund expense ratio, denoted as “net skill”: $D(s_{j,t}^{BG}, \eta) = \tilde{a}_{j,t} - p_{j,t}$. So, with a log specification for $D(\cdot)$, we have

$$\log(s_{j,t}^{BG}) = \frac{\tilde{a}_{j,t} - p_{j,t}}{\eta}. \quad (2.16)$$

This expression is intuitive: the higher the net skill of a fund, the larger is the efficient fund size; the stronger the effect of decreasing returns to scale, the smaller the fund’s size will be.

To compare BG-implied fund size with data, we construct ten portfolios of mutual funds sorted on net skill. We then compute mean of log size in the data and in the BG model.
for each portfolio.\textsuperscript{16} Figure 2.1 presents the result. First, we can see that in the data, the mean fund size monotonically increases with net skill. This result is consistent with the Berk and Green model’s prediction. But we also witness a discrepancy between the data and the model. On the higher end, BG predicts the mean size of funds in portfolio 10 to be 7.3 billion. In the data, the mean fund size in portfolio 10 is 936 million. On the lower end, according to BG, the mean fund size in portfolio 1 is 0.7 million. And in the data, it is 134 million. These differences are statistically significant as indicated by the 95-percent confidence intervals. From this figure, we can draw the conclusion that asset misallocation exists in both bad funds and good funds in the data.

[Insert Figure 2.1 Here]

The key prediction of Berk and Green (2004) is that asset inflows into funds that are estimated to be skilled based on their past returns will erode their subsequent performance due to decreasing returns to scale. In addition, fund managers who have been revealed as skilled raise their fees. As a result the net alpha of these funds should be zero in the future. We can test this prediction of the (generalized) model by looking at abnormal returns on the portfolios of mutual funds formed on their net skill discussed above. Thus, using the updated belief about fund skill as well as its fees in year $t$, funds are placed into portfolios, and we track equal-weighted returns on these portfolios over the subsequent 12 months, until the portfolios are re-sorted based on the updated information from $t+1$. Estimated five-factor alphas on these portfolios along with their 95% confidence intervals are displayed in Figure 2.2. Consistent with much of the mutual fund literature, the vast majority of alphas are negative (i.e., for all but the top two deciles of net skill). Perhaps more surprisingly, funds in the top decile of estimated net skill actually do display statistically significant outperformance. Overall, realized alpha is monotonically increasing with the estimated net skill, ranging from close to $-3\%$ per annum for the funds in the bottom decile, to about $0.7\%$ for those in the top decile. This result indicates a stark rejection of the key prediction of

\textsuperscript{16}We winsorize the belief $\tilde{a}$ at 1\% and 99\% level because there are some outliers in the estimated beliefs.
the BG model. Not only don’t assets seem to flow out of unskilled and/or expensive funds towards the relatively more skilled ones, as suggested by evidence in Figure 2.1 above, but even the “best” funds that are “too small” relative to the BG model don’t raise their prices sufficiently to fully capture their outperformance. Thus the observed allocation of capital across equity mutual funds, combined with their price setting behavior, present a quantitative puzzle for the frictionless BG model.

[Insert Figure 2.2 Here]

2.5.3. Search model parameters

Our search model is meant to bridge the gap between the efficient capital allocation described by the Berk and Green model and the actual allocation of assets across mutual funds observed in the U.S. data. Table 2.2 reports the estimated parameters of the structural search model. With the view towards conducting counterfactual analysis, we rely on the more recent sample of the data for this part of the estimation, choosing 2001 as our starting point (our estimation results are robust to various starting points, however). As described in the estimation section, we estimate the model using four versions of the moment conditions in (2.15). All the parameters other than $\theta$ are quite stable across the four sets of estimates. This assures us that even though our identification of $\theta$ relies on one subsample, it does not affect other parameters drastically.

[Insert Table 2.2 Here]

Our estimate of $\lambda$, the mean of search cost, is 39 basis points. Hortasu and Syverson (2004) find that the mean search cost for the S&P 500 index fund market is from 11 bps to 20 bps across different specifications. Our estimated average search cost is somewhat higher than theirs since, presumably, investors in their sample have higher than average

\footnote{Hortasu and Syverson (2004) estimated two variants of their search model: one in which sampling probabilities are different across funds, and another one where they are identical. We view our model as being closer to the former. The estimation results for that version of the model are reported in Table III of their paper. The log mean search cost is around -6.17 to -6.78, which implies mean search costs ranging from 11 bps to 20 bps.}
level of financial sophistication (implying a lower level of search costs). This is because they focus on investment in S&P 500 index funds in the late 90s, when these funds were not as prominent as they are today. Alternatively, it may be the case that it is harder to evaluate actively managed mutual funds (compared to index funds, which are relatively simple products), and hence implied barriers to information acquisition that are implied by the observed distribution of fund size are greater.

The magnitude of the mean search cost is quite significant. For the average investor, the cost of drawing another sample fund is 39 basis points, which is comparable to the mean alpha in our sample. The large magnitude of estimated search cost is a reflection of the active fund under-performance puzzle. In the mutual fund literature, numerous papers documented the (persistent) underperformance of (at least a large subset of) active funds (e.g., Carhart (1997)). Since many under-performing funds enjoy sizable market shares, our model requires a high search cost to rationalize those facts. In our model, high search cost investors will find it suboptimal to continue searching for a better fund than those drawn in the first couple of attempts. In the counterfactual case, if the search costs were low then index funds would be much larger than observed in the data, and underperforming active funds would be substantially smaller.

Our key parameter of interest is $\theta$, the coefficient in front of marketing expenses in the sampling probability function. First, we notice that the estimated $\theta$ is the smallest when we use the moment conditions of the funds that choose to do no marketing, and the largest for the funds that choose the upper bound of 1%. For the funds that choose the interior levels, $\theta$ is in the middle. This is intuitive because $\theta$ measures the effectiveness of marketing. The funds that are at the upper bound are more likely to be constrained in their ability to increase their marketing in an effort to increase investors’ awareness. Consequently, the first order condition (2.15) is likely to be satisfied with an inequality, and forcing it to be zero in estimation biases the estimate upward. Similarly, the funds at the lower bound are likely to find it optimal to receive a “rebate” on marketing in order to increase their profits,
but since such rebates are not available their first order condition is also likely not to be satisfied, biasing the estimate of $\theta$ downwards. In what follows, we rely on the estimates obtained with funds in the interior of the marketing expenditures as our benchmark.

In the sampling probability function, besides marketing expenses, we include fund family size, log fund age and year fixed effect. The coefficient of family size is positive and significant, confirming the idea that larger fund families are better at informing investors about their products. The fund age coefficient is positive and significant, which is intuitive, as older funds also have more visibility than younger funds. This result is also consistent with Hortasu and Syverson (2004) evidence from the S&P 500 index fund market.

In order to put these estimates into perspective, we conduct the following experiments. We compute the percentage changes in fund size for various groups of funds when marketing expense increases by 1 bp. Table 2.3 provides the results. Each column corresponds to results computed using different values of the $\theta$ parameter - those obtained with the funds on the upper bound ($\theta = 133.18$), lower bound ($\theta = 111.22$), and in the interior ($\theta = 113.11$) of marketing expenditures, as described above. All the other parameters are fixed at the benchmark levels (estimates from the “interior” funds). When we change fund $j$’s marketing, we fix all the other funds’ prices and marketing expenses and fund $j$’s price (i.e., this is a comparative static, not counter-factual analysis). Thus, holding total fees fixed, a 1 bp increase in marketing implies an equivalent reduction in profit margins.

[Insert Table 2.3 Here]

Overall, a 1 bp increase in marketing expenses leads to a roughly 1% increase in fund’s size, but there is substantial variation across different types of funds, and this elasticity naturally increases with $\theta$. In panel A, we sort funds by their size. We find that as fund size decreases, the sensitivity of size to a 1 bp increase in marketing rises. Using the benchmark estimates ($\theta = 113.11$) it goes from approximately 0.87% for large funds to 0.9% for small funds. This is intuitive because as a prior, marketing investment should be much more effective
for smaller funds because they have smaller probabilities of being known (e.g., typically, they are younger). Investing in marketing is a good way for small funds to attract greater investor attention. Interestingly, this sensitivity is higher both at the upper and at the lower estimates of $\theta$.

In panel B, we sort funds by their skill level $\tilde{\alpha}$. We find that marketing is much more useful for highly skilled funds. If high-skill funds can get into the consideration sets of more investors they will be picked by more investors. But for the low-skill funds, even if they are known to more investors, their size will not increase sufficiently to justify the extra expense. In fact, in Figure 2.3 we show that, for a fund of average age and belonging to a fund family of average size, with fund/year shock $\xi = 0$, the optimal level of marketing is increasing in the posterior belief about its skill. This result indicates that marketing is complementary to skill, yet it does not mean that it helps improve welfare in the presence of the search friction, since high-skill funds may be forced to spend “too much” on marketing, leading to a wasteful “arms race.”

[Insert Figure 2.3 Here]

Lastly, in panel C, we sort funds by their original marketing expenses levels. Lower Bound funds are funds that originally choose zero marketing expenses. Upper Bound funds are funds that originally chose 1% marketing expenses. Non binding funds are the rest of funds, which choose interior marketing levels. We find that an additional 1 bp increase in marketing is not very useful to funds at the upper bound (suggesting that many of these funds are at suboptimally high levels of marketing, perhaps due to inertia). Similarly, for funds at the lower bound extra marketing appears more worthwhile. Some of these funds might belong to fund families that choose to sell their funds directly rather than through brokers, for example, and as a consequence do not charge any 12b-1 fees, even though it might be beneficial for some of their funds.

Next we analyze the impact of marketing on fund profits. Table 2.4 displays the results. In
panel A, we sort funds by size; we find that for the small funds extra marketing increases profits, if all the other funds’ strategies in pricing and marketing stay the same (since we are not recomputing their best responses in this exercise). In panel B, we show that when \( \theta \) is at the higher level of estimates, it is profitable for high-skill funds to do more marketing. In panel C, we find that essentially all of the funds are worse off if they increase their marketing, which is not surprising given that the estimation procedure assumes that funds are at their optimal levels of marketing (on average).

[Insert Table 2.4 Here]

2.5.4. Sampling probabilities and fund size

In this section, we quantify the impact of various components of the sampling probability on explaining the size distribution of funds. Our method is as follows: we first set a particular component in the sampling probability equation (2.6) to equal zero. For example, by setting all funds’ marketing expenses \( b \) equal zero, we are effectively removing all the explanatory power of marketing expenses from the model. Then we recompute the model-implied market shares of all funds using equation (2.9), which captures investor demand side of our model. Specifically, let

\[
s^*_t = F_t [p^*_t, b^*_t, a^*_t - \eta \log (M_t s_t), x^*_t, \xi^*_t, p_{0,t}; \Theta],
\]  

(2.17)

where the the arguments with the asterisks are equal to their empirical values if the variable is “included” in the specification, and otherwise set to zero (for fund skill \( \tilde{a}_j \) and expense ratio \( p_j \) we use sample means when the variables are “not included” in the specification.\(^{18}\) Importantly, in this exercise we are not recomputing the whole equilibrium, since we keep other variables fixed (rather than solving for every fund’s best response).

We regress the log of market share of funds in the data \( s_{jt} \) on these reduced model-implied log market shares \( s^*_j,t \) and report the R-squared of these regressions in Table 2.5. The lower the R-squared, the more important that component is in terms of explaining the size

\(^{18}\)An explicit recursive expression for the market share function \( F \) is provided in the Appendix.
distribution. Among all of them, the unobserved characteristics of the fund $\xi$ is the most important one (responsible for almost half of the R-squared). This is reasonable because we only include a limited number of variables in our estimation; any other variables that could potentially affect fund size would be subsumed by $\xi$. The second most important variable is age. After controlling for fund’s age and other variables, the family size doesn’t add much explanatory power.

What about the key features of mutual funds that have been the main focus of the literature - skill and costs? Removing either variation in posterior skill or in the fund price (expense ratio) reduces the R-squared to about 90% in each case. Importantly, removing instead the marketing variable yields a very similar R-squared of 92%. This indicates that marketing is nearly as important in terms of explaining the size distribution of mutual funds as price or skill.

We are also interested in understanding how do various components of our model contribute to the misallocation of capital to funds. We compute the correlation between the reduced model-implied fund size $s_{j,t}^*$ and the BG-implied fund size $s_{j,t}^{BG}$, as defined in equation (2.16). We can see that in the data (or, equivalently, our unrestricted model, which matches the data by construction) the correlation is positive but small, at 0.09. If changing one of the components of the model increases this correlation, that means that this change makes capital allocation more efficient. This correlation is at its highest level of 0.59 if we only include fund skill and price. Conversely, removing price or skill but including other (search model) ingredients reduces this correlation, since these are the key elements of the Berk and Green model. At the same time, removing marketing increases the correlation. This means that marketing could potentially account for at least some of the misallocation that we observe in the data. However, this analysis is only suggestive, since we do not compute the optimal response of the funds to the induced change. In what follows, we describe counterfactual experiments that fully take into account the equilibrium behavior of both investors and funds.
2.6. Counterfactual Analysis

Section 2.5.2 documents substantial capital misallocation in the mutual fund industry. In this section, we use our model to quantitatively study the importance of marketing expenses and search costs in shaping the equilibrium fund size and expense ratios. We also investigate how they affect allocational efficiency and investor welfare. First, we explore a counterfactual equilibrium with no marketing.\textsuperscript{19} We then investigate the impact of changing search costs on equilibrium marketing expenses. We focus on the most recent year in our sample (2015) for these experiments.

2.6.1. Welfare measures

We first present how the welfare of different parties in the market are calculated. Fix a year $t$ (the time subscript $t$ will be suppressed in this section). In our model, investor’s utility consists of two parts, the expected indirect utility provided by the fund that investor chooses and the expected total search costs the investor incurs in order to find this fund. The welfare of investor $i$ with search cost $c_i$ is given by

$$V(c_i) = \int_{\bar{u}(c_i)}^{+\infty} u d\bar{u}(u) \Psi(\bar{u}) \frac{\Psi[\bar{u}(c_i)]}{1 - \Psi[\bar{u}(c_i)]} - c_i \Psi[\bar{u}(c_i)] \frac{1}{1 - \Psi[\bar{u}(c_i)]},$$

(2.18)

where $\bar{u}$ is the reservation level of indirect utility (detailed derivation of investor’s welfare is provided in the appendix). For a higher level of reservation utility, the investor needs to search more in order to find the desired fund. We see that the expected total search cost $c_i \frac{\Psi[\bar{u}(c_i)]}{1 - \Psi[\bar{u}(c_i)]}$ is increasing in $\bar{u}$. In the first term of equation (2.18), the numerator is the expected indirect utility for the funds with higher than $\bar{u}$ utility level. The denominator adjusts for the fact that the investor will only pick the funds from this part of the distri-

\textsuperscript{19}Recently the SEC considered a proposal to improve the regulation of mutual fund distribution fees, in particular, by limiting fund sales charges as a way of protecting retail consumers from unnecessarily high costs. Our counterfactual analysis can be viewed as analyzing welfare consequences of a policy that set the marketing cap at zero.
bution. The aggregate measure of investor welfare in this model is derived by integrating across the search cost distribution:

\[ U = \int_0^{+\infty} V(c_i) dG(c_i). \]  

(2.19)

Fund profits is also part of the total welfare. These include the profits for both active funds and index funds:

\[ P = \sum_{j=1}^{N} (p_j - b_j)s_j + s_0p_0. \]  

(2.20)

Here the first part is the total profits for the active funds, the second part is the total profits for the passive funds. In the counterfactual analysis, we assume index fund price is fixed and we resolve the equilibrium for the active funds’ prices and marketing expenses. In our counterfactual we assume $M$ stay the same.

If marketing expenses constitute pure payments to labor (e.g., broker commissions) rather than dead weight costs, they should also be considered in the welfare analysis:

\[ B = \sum_{j=1}^{N} b_j s_j \]  

(2.21)

Our measure of total welfare is the sum of the three components above: $U + P + B$.

2.6.2. Equilibrium with no marketing

In this simulation, we restrict marketing expenses to zero. We use year 2015’s data and the benchmark parameters from column (1) in Table 2.2. Table 2.6 provides the comparison between the currently observed equilibrium and the no-marketing equilibrium on some of the key measures. First, the mean expense ratio drops by almost 77 basis points in the counterfactual relative to the current equilibrium. This drop is larger than the decrease in the average marketing expenditure. It indicates fiercer price competition between funds.
when they cannot attract investors through marketing. To further understand the price changes across funds, we split the funds into four groups based on their marketing expenses in the current equilibrium: (1) funds whose marketing is at the upper bound of 100 bps, (2) funds whose marketing is at the lower bound of 0, (3) funds whose marketing is between 1 bp to 49 bps and (4) funds whose marketing is between 50 bps and 99 bps. We plot the price differences for all funds between the current equilibrium and the no-marketing equilibrium in Figure 2.4 panel A. We find that all the funds lower their prices in the no-marketing equilibrium, but the magnitude of change varies substantially across the four groups. Group (1) funds lower their prices by around 100 bps (i.e., roughly their original marketing costs). The most interesting finding is that the group (2) funds in the no-marketing equilibrium also lower their prices by around 30 bps, which necessarily has to come from a reduction in profit margins since these funds do not have any marketing expenses in the current equilibrium. This is mainly due to the effect of competition between funds. A similar but weaker effect is present for most of the funds in groups (2) and (3).

Second, we find that the total market share of active funds drops from 74% to 68%. This indicates that marketing is useful for steering investors towards active funds. When funds cannot do marketing, they lose market share since they are less likely to enter investors’ information sets. The sampling probability of index funds increases. This is due to the assumption that all the sampling probabilities sum to 1. When active funds cannot do marketing, the index funds are more likely to be “found.” In the no-marketing equilibrium, active funds’ profits drop by 15 basis points on average. This is resulting from both the shrinking of the total market share and the fall of profit margins.

Investor welfare in the no-marketing equilibrium increases by around 57%. There are three main contributing factors: lower prices, higher alphas, and lower search costs. As assets under management decline, the average alpha of the industry increases from 37 bps to 41 bps, due to the effect of decreasing returns to scale. In Figure 2.4 panel B, we plot the
difference in fund alphas between the no-marketing equilibrium and current equilibrium for different groups of funds. We find that for funds in group (1) alpha increases consistently. This is mainly because in the no-marketing equilibrium their assets under management fall and so, due to decreasing returns to scale, their alphas increase. For other groups of funds, some of the alphas increase, while others decrease.

We can compute the aggregate search cost incurred by the investors in the two equilibrium. The aggregate search cost is given by:

\[ \int_{0}^{+\infty} c_i \frac{\Psi[\bar{u}(c_i)]}{1 - \Psi[\bar{u}(c_i)]} dG(c_i). \] (2.22)

We find aggregate search costs are lower in the no-marketing equilibrium. In the model, investors search until the expected benefit of finding better funds is smaller than the unit search cost. If investor \( i \) has already found fund \( j \) with utility \( u_j \), then his incentive to search hinges on both her search cost \( c_i \) and the expected possible gain from continuing the search. If there are not too many better funds out there, then investor’s incentive to search is weaker. To show that this is indeed the reason why investors search less in the no-marketing equilibrium, we plot the histogram of indirect utilities associated with individual funds in the two equilibria in Figure 2.5. We find the standard deviation of utility levels in the no-marketing equilibrium is substantially lower than in the current equilibrium. Since the dispersion in available utilities is reduced, so are the expected benefits of searching and, consequently, investors search less. Through the resulting reduction in search costs, investor welfare increases by 17 basis points on average.

When marketing is eliminated the size of active funds doesn’t drop drastically for two reasons. First, there are characteristics in the sampling probability function, besides marketing expenses, which ensure that all active funds sampling probabilities are positive. Second,
lowering the size of active funds increases their performance. This effect makes active funds more attractive.

2.6.2.1. Heterogeneous effect across investors

In our model, we assume different investors have different search costs. In this section, we study the impact of eliminating marketing across investors with different levels of search costs. We focus on the following dimensions for each investor: individual investor welfare, total incurred search cost, gross alpha expected by investors, total expense ratios investors pay, and marketing expenses that investors implicitly pay for as part of their chosen funds’ expense ratios (in expectation). Figure 2.6 panel A shows that for all the search cost levels, in the no-marketing equilibrium, investors achieve a higher level of welfare on average. But the biggest improvements come from the high search cost investors. Their welfare increases roughly by 100 basis points. For the low search cost investors the increase is not very large. This is because the low search cost investors always find the “best” funds available in the market. Figure 2.6 panel B shows a somewhat non-monotonic relationship between unit search cost and total search costs incurred. For the low search cost investors the total search cost is not very high even though they search a lot since their unit search costs are low. The high search cost investors find it too costly to conduct any search, so they search infrequently, many stopping after the first (free) search. Consequently, high search cost investors’ total search cost is also low. The intermediate search cost investors search relatively aggressively and their search costs are non-trivial. So in total they incur the largest total search costs. Comparing the two equilibria, we find that in the no-marketing equilibrium, the intermediate search cost investors incur lower total search costs. This is due to the fact that in the no-marketing equilibrium, average fund quality improves, so that it is easier for investors to find funds that satisfy their reservation levels.

Focusing on Figure 2.6 panel C and panel D, we find that in general, high search cost investors get lower alpha funds, pay high prices and high marketing expenses. This is simply because high search cost investors don’t search very much. An interesting fact is
that for the very low search investors, the funds they invest in have positive net alphas. In Berk and Green’s model, since investors have zero search cost, in equilibrium, all the funds have zero net alphas. But in our model, since all investors incur a positive search cost, the low search investors are able to find funds that are both skilled and cheap, but not found by enough other investors, so that their performance is not fully eroded by decreasing returns to scale. At the same time, these high-skill funds do not find it optimal to increase their expense ratios (and thus drive net alphas towards zero) because that would make these funds less attractive to the more discerning (low search cost) investors, whose choices are very sensitive to fees.

2.6.3. Allocational efficiency

It is also interesting to consider the consequences of restricting marketing on capital allocation within the mutual fund sector. On the one hand, we see that average (gross) fund alpha increases in the no-marketing equilibrium, suggesting that some highly skilled funds might be “too small,” operating below their efficient scale. Indeed, since we show that in the current equilibrium highly skilled funds benefit more from marketing, ceteris paribus, it is reasonable to expect that without the ability to do any marketing these funds might be disproportionately hurt by the imposed constraint. On the other hand, marketing is an important driver of costs, which are in turn a major determinant of net alphas (and indirect utilities) enjoyed by investors.

In keeping with our initial approach, we compare fund size distribution implied by the frictionless benchmark in the style of Berk and Green (2004) and that generated by our search model counterfactual. Figure 2.7 provides the comparison for the year 2015. Panel A displays the direct analogue of Figure 2.1 restricted to the data for 2015: the BG-implied values are computed using the posterior beliefs about fund alphas as well as their observed expense ratios and the estimated decreasing returns to scale parameter (the fund size in the
data is consistent with the search model by construction). Panel B presents the analogues of these values in the counterfactual equilibrium with no marketing. That is, the “counterfactual” plot uses the fund size computed under the counterfactual equilibrium, where as the “BG-implied” values are recomputed using the expense ratios in the counterfactual equilibrium.

We observe that in the no-marketing case the two lines are much closer to each other than in the current equilibrium. This is true only in small part due to the steepening in the relationship between log size and net skill, visible mostly in the middle of the skill distribution. The changing BG-implied distribution plays a noticeably more important effect. This is due to the fact that funds are charging substantially lower fees to their investors in the no-marketing equilibrium. Thus the solid black line in the graph shifts upward, closer to the blue line. The shift appears especially pronounced for the lowest-skill funds, even though they are still “too big” in the counterfactual relative to the frictionless model, where as for the highest-skill funds there is not much difference between the two measures. Thus, the overall effect of eliminating marketing expenditures is to improve the efficiency of capital allocation in the active fund industry, at least from the standpoint of net abnormal returns to investors as emphasized in Berk and Green (2004).

2.6.4. Reducing search costs

Last, we examine the impact of search costs on equilibrium market outcomes with special attention to marketing expenses. Because of search costs, competing on marketing could be a potential profitable strategy for some funds, since they essentially just need to be sampled by the least-discerning high-cost investors frequently enough. But with the emergence of the Internet, advancement in search technologies (e.g., Google), more transparent comparison (e.g., services like Morningstar and Lipper), and better investor education, we would expect the search frictions to decline over time. In order to analyze the potential impact of new technologies we consider a counterfactual equilibrium where we set the mean search cost to 35 bps or 20 bps. Given the new search cost, funds reoptimize their prices and marketing
expenses. We find that as the average search cost decreases from 39 bps to 35 bps, mean marketing expenses drop from 61 bps to 44 bps. But when the mean search cost further drops to 20 bps, the equilibrium marketing expenses become zero. Notice that the regulatory cap is still held at 100 bps. The intuition is as follows: low search costs render marketing less profitable. In the model with high mean search cost, there exists a large fraction of investors with very high search costs. A subset of funds specifically exploit these “unsophisticated” investors. Those funds invest aggressively in marketing so as to enter more of the high search cost investors’ choice sets. Since such investors will not search much, they do end up investing with those funds even if they are not very skilled and quite expensive. But when mean search cost drops to sufficiently low level, this strategy is no longer profitable, since the model presumes there are fewer investors who find it too costly to continue searching for a better fund. Therefore, when search costs are not very high, funds will not invest in marketing and instead compete on price. This result provides a new perspective on the recent evolution of the asset management industry documented by Stambaugh (2014): declining fees charged by active funds coincident with the growth in passive index funds. From the standpoint of our model, both trends can be seen as resulting from falling search costs, due to a combination of information technology and growing investor sophistication.

2.7. Concluding Remarks

The question whether actively-managed mutual funds exhibit skill - i.e., persistent outperformance - has a long history in financial economics, since it is central to the debate about informational efficiency of securities markets in the sense of Fama. While there is still substantial debate about the ability of an “average” fund manager to generate abnormal returns (before or after fees are taken into account), perhaps one of the most robust findings in the literature is that investors’ flows are much less sensitive to past bad performance than to outperformance (Ippolito (1992), Carhart (1997), Chevalier and Ellison (1997), Sirri and Tufano (1998b), etc.). This evidence hints that the market for mutual funds may not be
efficient at allocating capital across funds because bad funds aren’t punished sufficiently for poor performance, and therefore underperforming managers control more assets than justified by their level of skill. Capital misallocation in the mutual fund industry could potentially lead to inefficiencies in capital allocation across firms, distorting real investment (van Binsbergen and Opp (2016)). It is therefore important to understand quantitatively how much capital is misallocated in the mutual fund industry. By estimating the Berk and Green model, we find that in the the U.S. equity mutual funds data, from year 1964 to year 2015, all but the best-performing decile of mutual funds are “too large” relative to the optimal scale predicted by the BG model. These results indicate that there exist substantial frictions in the market for mutual funds.

In our paper, we view mutual fund marketing expenses as purely informative (e.g., Butters (1977)). It is possible that a portion of these marketing expenses serves a persuasive function in ways highlighted in the theoretical literature: e.g., firms may find it profitable to steer investors toward non-price attributes (Mullainathan, Schwartzstein, and Shleifer (2008), Gabaix and Laibson (2006), Carlin (2009), Ellison and Ellison (2009)). Separating the informative effect from the persuasive effect of marketing one would require information investors’ actual choice sets, which is generally not available. Thus, by making the assumption that all marketing is informative, our welfare analysis results provide an upper bound on the social value of mutual fund marketing.\textsuperscript{20} Relaxing this assumption in order to understand the possible welfare loss from “persuasive” marketing is a fruitful venue for future research.

\textsuperscript{20}Even if marketing is purely informative, due to the externality of marketing in our model, marketing investment can still be excessive. Fund $i$’s marketing investment could decrease fund $j$’s probability of being known. In a Nash equilibrium, funds will not take the externality into consideration when deciding the marketing investment levels. All of the funds might be better off if they agree on a lower level of marketing investment - But of course, this agreement is fragile since deviation is profitable.
This figure plots the mean of log fund size (measured in millions of dollars) for portfolios of funds formed on net skill (defined as posterior belief about the fundamental skill level $\tilde{a}$ minus expense ratio $p$). We compute fund size according to the generalized version of the Berk and Green (2004) model that we estimate: $\log(s_{j,t}^{BG}) = \frac{\tilde{a}_{j,t} - p_{j,t}}{\eta}$, where $\eta$ captures decreasing returns to scale. The black line plots the mean of the Berk and Green model-implied fund sizes for each portfolio (BG). The blue line plots the mean of log fund size in the data for each portfolio. Portfolio 1 has the lowest net skill while portfolio 10 has the highest net skill. 95 percentile confidence bounds are indicated by dashed lines.
Figure 2.2: Capital (mis)Allocation in Mutual Funds: Net Alpha vs. Net Skill

This figure plots the average annual net alpha for portfolios of funds formed on net skill (defined as posterior belief about the fundamental skill level $\tilde{a}$ minus expense ratio $p$). The black line plots the Berk and Green (2004) model-implied net alpha for each portfolio (BG). The blue line plots the mean of net alpha in the data for each portfolio. Portfolio 1 has the lowest net skill while portfolio 10 has the highest net skill. 95 percentile confidence bounds are indicated by dashed lines.

Figure 2.3: Marketing and Skill

This figure plots the relationship between fund’s skill $\tilde{a}$ and model-implied fund’s marketing expense, for a fund in year 2015 that has average characteristics, $\xi = 0$, $\zeta = 0$, and $\omega = 0$. We vary the posterior belief about its skill $\tilde{a}$ and calculate the associated optimal marketing expense, given the choices of the other funds observed in the data.
Panel A plots the expense ratio reduction as a result of moving from the current equilibrium (which allows marketing) to the counterfactual no-marketing equilibrium. The x-axis is fund size (here we use log market share). The y-axis is the current equilibrium price minus the no-marketing equilibrium price.

Panel B plots the differences in gross alpha between the current equilibrium and the no-marketing equilibrium. The x-axis is fund size (here we use log market share). The y-axis is the gross alpha changes.

We split funds into 4 groups based on their current marketing expenses. Group 1, indicated by squares, marketing expenses of 100 bps. Group 2, indicated by asterisk, marketing expenses of 0 bp. Group 3, indicated by diamond, marketing expenses between 1 and 49 bps. Group 4, indicated by circle, marketing expenses between 50 and 99 bps.
Figure 2.5: Indirect Utilities: Current Equilibrium vs. No-Marketing Equilibrium

The left panel shows the histogram of indirect utilities associated with funds in the no-marketing equilibrium. The right panel shows the histogram of indirect utilities associated with funds in current equilibrium (which allows marketing). The x-axis is utility level. The y-axis is frequency. Indirect utility is defined in equation (2.5).
Panel A plots investor welfare as a function of unit search cost levels. The x-axis is investor’s unit search cost, \( c_i \) in basis points. The investor welfare is in unit of bp. The y-axis is investor welfare defined as indirect utility provided by chosen fund minus total incurred search cost. For the expression of investor welfare as a function of search cost, please refer to equation (2.18).

Panel B plots investor’s expected total search cost as a function of unit search cost levels. Expected total search cost is defined as \( c_i \Psi[u(c_i)] - \Psi[\bar{u}(c_i)] \), where \( c_i \) is the search cost level, \( \Psi[u(c_i)] \) is the probability of sampling a fund that delivers the investor an indirect utility smaller or equal to \( \bar{u}(c_i) \) (detailed derivation is in appendix.)

Panel C plots the gross alpha expected by investors as a function of unit search cost levels.

Panel D plots the expense ratios and marketing expenses investors incur as a function of the unit search cost levels.

Solid line stands for the current equilibrium and dashed line stands for no-marketing equilibrium.
Figure 2.7: Capital (mis)Allocation: Counterfactual

Panel A plots the mean of log fund size (measured in millions of dollars) for portfolios of funds formed on net skill (defined as posterior belief about the fundamental skill level $\hat{a}$ minus expense ratio $p$) for the current equilibrium, using data on mutual funds in the year 2015. Panel B plots the mean of log fund size (measured in millions of dollars) for portfolios of funds formed on net skill for the no-marketing equilibrium.

We compute fund size according to the generalized version of the Berk and Green (2004) model that we estimate: $\log(s_{BG,j,t}) = \hat{a}_{j,t} - p_{j,t} \eta$, where $\eta$ captures decreasing returns to scale.

The black line plots the mean of log fund size for each portfolio implied by the Berk and Green (2004) model. The blue line plots the mean of log fund size in the data or in the counterfactual for each portfolio. We construct ten portfolios based on the deciles of net skill. Portfolio 1 has the lowest net skill while portfolio 10 has the highest net skill.
Table 2.1: Investor Beliefs and Manager Skill Parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Description</th>
<th>1964-2015</th>
</tr>
</thead>
<tbody>
<tr>
<td>η</td>
<td>Decreasing returns to scale (%)</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.04)</td>
</tr>
<tr>
<td>μ</td>
<td>Mean of prior (%)</td>
<td>3.05</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.25)</td>
</tr>
<tr>
<td>κ</td>
<td>SD of prior (%)</td>
<td>2.41</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.12)</td>
</tr>
<tr>
<td>δ</td>
<td>SD of realized alpha (%)</td>
<td>7.62</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.05)</td>
</tr>
<tr>
<td>ρ</td>
<td>Skill persistence</td>
<td>0.94</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.02)</td>
</tr>
<tr>
<td>Num of Obs</td>
<td></td>
<td>27,621</td>
</tr>
</tbody>
</table>

This table presents the estimates of the fund performance-related parameters. The decreasing returns to scale parameter, η, the mean of manager’s *ex ante* skill distribution, μ, the standard deviation of this distribution, κ, the standard deviation of the idiosyncratic noise in the realized alpha, δ, and the persistence of the manager’s skill, ρ. The standard errors are in the parentheses.

Table 2.2: Search Model Parameters

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Description</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>λ</td>
<td>Mean search cost (bp)</td>
<td>39</td>
<td>39</td>
<td>39</td>
<td>39</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(4)</td>
<td>(4)</td>
<td>(4)</td>
<td>(4)</td>
</tr>
<tr>
<td>γ</td>
<td>Alpha weight</td>
<td>0.415</td>
<td>0.415</td>
<td>0.416</td>
<td>0.415</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.030)</td>
<td>(0.031)</td>
<td>(0.030)</td>
<td>(0.030)</td>
</tr>
<tr>
<td>θ</td>
<td>Marketing effectiveness</td>
<td>113.11</td>
<td>111.22</td>
<td>133.18</td>
<td>122.56</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(7.334)</td>
<td>(7.291)</td>
<td>(8.797)</td>
<td>(7.397)</td>
</tr>
<tr>
<td>β₁</td>
<td>Family size effect</td>
<td>0.4048</td>
<td>0.403</td>
<td>0.381</td>
<td>0.393</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.026)</td>
<td>(0.026)</td>
<td>(0.026)</td>
<td>(0.026)</td>
</tr>
<tr>
<td>β₂</td>
<td>Fund age effect</td>
<td>1.032</td>
<td>1.032</td>
<td>1.032</td>
<td>1.032</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.037)</td>
<td>(0.038)</td>
<td>(0.037)</td>
<td>(0.036)</td>
</tr>
<tr>
<td>Year FE</td>
<td></td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

This table presents the estimates of the structural search model. We use the data from 2001 to 2015. The four columns correspond to four sets of moment conditions as described in Section 2.4. In column (1), we use the funds that are not constrained in their marketing expenses to estimate the model. In column (2), we use only funds whose marketing expenses are 0 to estimate the model. In column (3), we use the funds whose marketing expenses are 100 bps to estimate the model. In column (4), we use all of the funds to estimate the model.
Table 2.3: Change in Size when Marketing Expenses Increase by 1 bp

<table>
<thead>
<tr>
<th></th>
<th>$\theta = 113.11$</th>
<th>111.22</th>
<th>133.18</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Sort by Size</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Big Funds</td>
<td>0.8735</td>
<td>0.8904</td>
<td>1.043</td>
</tr>
<tr>
<td>Intermediate Size Funds</td>
<td>0.8794</td>
<td>0.8965</td>
<td>1.050</td>
</tr>
<tr>
<td>Small Funds</td>
<td>0.9085</td>
<td>0.9261</td>
<td>1.085</td>
</tr>
<tr>
<td><strong>Panel B: Sort by Skill</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High Skill Funds</td>
<td>0.9670</td>
<td>0.9858</td>
<td>1.155</td>
</tr>
<tr>
<td>Intermediate Skill Funds</td>
<td>0.8987</td>
<td>0.9161</td>
<td>1.073</td>
</tr>
<tr>
<td>Low Skill Funds</td>
<td>0.8154</td>
<td>0.8311</td>
<td>0.973</td>
</tr>
<tr>
<td><strong>Panel C: Sort by Original Marketing Expense</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Binding at Lower Bound</td>
<td>0.9554</td>
<td>0.9739</td>
<td>1.141</td>
</tr>
<tr>
<td>Non-Binding</td>
<td>0.8990</td>
<td>0.9165</td>
<td>1.073</td>
</tr>
<tr>
<td>Binding at Upper Bound</td>
<td>0.8413</td>
<td>0.8575</td>
<td>1.004</td>
</tr>
</tbody>
</table>

This table provides the percentage changes in fund size resulting from a 1 basis point increase in marketing expenses for various groups of funds using parameters from Table 2.2 column (1) except $\theta$. For $\theta$, we use the estimated value of $\theta$ from columns (2), (1) and (3) respectively in Table 2.2. In panel A, we sort funds by size. “Big Funds” are funds in the top decile. “Small Funds” are funds in the bottom decile. “Intermediate Size Funds” are the rest. In panel B, we sort funds by skill level. “High Skill Funds” are funds in the top decile. “Low Skill Funds” are funds in the bottom decile. “Intermediate Skill Funds” are the rest. In panel C, we sort funds by marketing expenses. “Binding at Lower Bound” are funds with no marketing expenses. “Binding at Upper Bound” are funds whose marketing are at the upper bound of 1%. “Non-Binding” are the rest.
Table 2.4: Change in Profits when Marketing Expenses Increase by 1 bp

<table>
<thead>
<tr>
<th>Panel A: Sort by Size</th>
<th>( \theta = 113.11 )</th>
<th>111.22</th>
<th>133.18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Big Funds</td>
<td>-0.4317</td>
<td>-0.4150</td>
<td>-0.2645</td>
</tr>
<tr>
<td>Intermediate Size</td>
<td>-0.0311</td>
<td>-0.0143</td>
<td>0.1381</td>
</tr>
<tr>
<td>Small Funds</td>
<td>0.0850</td>
<td>0.1024</td>
<td>0.2602</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B: Sort by Skill</th>
<th>( \theta = 113.11 )</th>
<th>111.22</th>
<th>133.18</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Skill Funds</td>
<td>-0.1267</td>
<td>-0.1081</td>
<td>0.0597</td>
</tr>
<tr>
<td>Intermediate Skill</td>
<td>-0.3358</td>
<td>-0.3186</td>
<td>-0.1634</td>
</tr>
<tr>
<td>Low Skill Funds</td>
<td>-0.2128</td>
<td>-0.1972</td>
<td>-0.0567</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel C: Sort by Original Marketing Expense</th>
<th>( \theta = 113.11 )</th>
<th>111.22</th>
<th>133.18</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binding at Lower Bound</td>
<td>-0.2100</td>
<td>-0.1917</td>
<td>-0.0263</td>
</tr>
<tr>
<td>Non-Binding</td>
<td>-0.1722</td>
<td>-0.1550</td>
<td>0.0006</td>
</tr>
<tr>
<td>Binding at Upper Bound</td>
<td>-0.4180</td>
<td>-0.4019</td>
<td>-0.2569</td>
</tr>
</tbody>
</table>

This table provides the percentage changes in fund profits resulting from a 1 basis point increase in marketing expenses for various groups of funds using parameters from Table 2.2 column (1) except \( \theta \). For \( \theta \), we use the estimated value of \( \theta \) from columns (2), (1) and (3) respectively in Table 2.2. In panel A, we sort funds by size. “Big Funds” are funds in the top decile. “Small Funds” are funds in the bottom decile. “Intermediate Size Funds” are the rest. In panel B, we sort funds by skill level. “High Skill Funds” are funds in the top decile. “Low Skill Funds” are funds in the bottom decile. “Intermediate Skill Funds” are the rest. In panel C, we sort funds by marketing expenses. “Binding at Lower Bound” are funds with no marketing expenses. “Binding at Upper Bound” are funds whose marketing are at the upper bound of 1%. “Non-Binding” are the rest.
Table 2.5: Quantifying the Importance of Sampling Probability Components

<table>
<thead>
<tr>
<th>Specification</th>
<th>ξ</th>
<th>age</th>
<th>num of family funds</th>
<th>marketing</th>
<th>skill</th>
<th>price</th>
<th>$R^2$</th>
<th>Correlation between $s_{BG}$ and $s_{Model}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>0.5169</td>
<td>0.2988</td>
</tr>
<tr>
<td>(2)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>0.2122</td>
<td>0.4988</td>
</tr>
<tr>
<td>(3)</td>
<td></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>0.1614</td>
<td>0.5698</td>
</tr>
<tr>
<td>(4)</td>
<td></td>
<td></td>
<td></td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>0.1152</td>
<td>0.5947</td>
</tr>
<tr>
<td>(5)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td>Y</td>
<td>Y</td>
<td>0.9157</td>
<td>0.1456</td>
</tr>
<tr>
<td>(6)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td></td>
<td>Y</td>
<td>0.9008</td>
<td>0.0301</td>
</tr>
<tr>
<td>(7)</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>0.9005</td>
<td>0.0776</td>
</tr>
</tbody>
</table>

This table presents results regarding quantifying the importance of various components in the sampling probability in explaining the size distribution of funds. Our method is as follows: we first set a particular component in the sampling probability equation (2.6) to equal zero. For example, by setting all funds’ marketing expenses $b$ equal zero, we are effectively removing all the explanatory power of marketing expenses from the model. Then, we recompute the model-implied market shares of all funds using equation (2.9) (we call them restricted model-implied fund size). Notice that here we are not recomputing the whole equilibrium. We fix all other variables and parameters. Lastly, we regress the log of market share of funds in the data onto model-implied log market shares and report the R-squared. We also report the correlation between the restricted model-implied fund size and the generalized Berk and Green model-implied fund size. “Y” indicates that we include this component in the sampling probability. Blank means that we remove this component. The data period is from 2001 to 2015.
Table 2.6: Summary of Outcomes for Current Equilibrium and No-Marketing Equilibrium

<table>
<thead>
<tr>
<th></th>
<th>Current</th>
<th>No-Marketing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean price (bp)</td>
<td>160.27</td>
<td>82.96</td>
</tr>
<tr>
<td>Mean marketing (bp)</td>
<td>61.29</td>
<td>0</td>
</tr>
<tr>
<td>Mean alpha (bp)</td>
<td>37.24</td>
<td>41.07</td>
</tr>
<tr>
<td>Total share of active funds</td>
<td>0.74</td>
<td>0.67</td>
</tr>
<tr>
<td>Mean sampling prob (%)</td>
<td>0.085</td>
<td>0.078</td>
</tr>
<tr>
<td>Sampling prob for low price funds (%)</td>
<td>0.042</td>
<td>0.14</td>
</tr>
<tr>
<td>Sampling prob for index fund (%)</td>
<td>5.91</td>
<td>13.66</td>
</tr>
<tr>
<td>Investor welfare (bp)</td>
<td>-140.72</td>
<td>-61.25</td>
</tr>
<tr>
<td>Active fund average profits (bp)</td>
<td>57.51</td>
<td>42.19</td>
</tr>
<tr>
<td>Index fund profits (bp)</td>
<td>2.32</td>
<td>2.86</td>
</tr>
<tr>
<td>Total welfare</td>
<td>-37.37</td>
<td>-16.20</td>
</tr>
<tr>
<td>Aggregate investor search cost (bp)</td>
<td>29.09</td>
<td>12.15</td>
</tr>
</tbody>
</table>

This table provides various measures of the mutual fund industry under current and no-marketing equilibrium. Mean price, mean marketing and mean alpha are the arithmetic average of price, marketing expenses and alpha for all active funds, respectively. Total share of active funds is the market share of all active funds. Sampling prob for low price funds is the mean sampling probability for the funds whose prices are below average. Investor welfare is defined in equation (2.19). Active fund average profits is the mean of price minus marketing expenses for all active funds. Index fund profits are defined similarly. Total welfare is the sum of investor welfare, funds’ total profits and total marketing expenses. Aggregate investor search cost is described in equation (2.22).
Table 2.7: Summary of Outcomes for Different Search Costs

<table>
<thead>
<tr>
<th></th>
<th>Low λ (20 bps)</th>
<th>Mid λ (35 bps)</th>
<th>High λ (39 bps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean price (bp)</td>
<td>58.52</td>
<td>136.24</td>
<td>160.27</td>
</tr>
<tr>
<td>Mean marketing (bp)</td>
<td>0</td>
<td>44.78</td>
<td>61.29</td>
</tr>
<tr>
<td>Mean alpha (bp)</td>
<td>38.94</td>
<td>39.00</td>
<td>37.24</td>
</tr>
<tr>
<td>Total share of active funds</td>
<td>0.64</td>
<td>0.71</td>
<td>0.74</td>
</tr>
<tr>
<td>Mean sampling prob (%)</td>
<td>0.07</td>
<td>0.08</td>
<td>0.08</td>
</tr>
<tr>
<td>Sampling prob for low price funds (%)</td>
<td>0.15</td>
<td>0.04</td>
<td>0.042</td>
</tr>
<tr>
<td>Sampling prob for index fund (%)</td>
<td>13.66</td>
<td>7.16</td>
<td>5.91</td>
</tr>
<tr>
<td>Investor welfare (bp)</td>
<td>-48.42</td>
<td>-118.41</td>
<td>-140.72</td>
</tr>
<tr>
<td>Active fund average profits (bp)</td>
<td>31.97</td>
<td>51.46</td>
<td>57.51</td>
</tr>
<tr>
<td>Index fund profits (bp)</td>
<td>3.15</td>
<td>2.58</td>
<td>2.32</td>
</tr>
<tr>
<td>Total welfare (bp)</td>
<td>-13.98</td>
<td>-33.04</td>
<td>-37.37</td>
</tr>
<tr>
<td>Aggregate investor search cost (bp)</td>
<td>9.18</td>
<td>25.75</td>
<td>29.09</td>
</tr>
</tbody>
</table>

This table presents various measures of the mutual fund industry under different search costs distributions. In the top row, there are three levels of mean search costs: 20bps, 35bps and 39bps. 39 bps is our estimated value from the data. Mean price, mean marketing and mean alpha are the arithmetic average of price, marketing expenses and alpha for all active funds, respectively. Total share of active funds is the market share of all active funds. Sampling prob for low price funds is the mean sampling probability for the funds whose prices are below average. Investor welfare is defined in equation (2.19). Active fund average profits is the mean of price minus marketing expenses for all active funds. Index fund profits are defined similarly. Total welfare is the sum of investor welfare, funds’ total profits and total marketing expenses. Aggregate investor search cost is described in equation (2.22).
A1. Appendix for Chapter 1

A1.1. Mutual Fund Manager’s Connections and Career Path

In this section, I investigate the relationship between the number of connections mutual fund managers have and their career outcome, with a special focus on termination, promotion, and demotion.

I follow Chevalier and Ellison (1999) by focusing on single-manager-managed funds. In my sample, there are 19,993 fund-year observations. Notice that this sample is not at the fund-manager-year level because one manager could manage more than one fund. To compute fund manager age, I require the manager’s birth year. I obtain birth years in the following ways: (1) Morningstar provides the birth year for a subset of fund managers or (2) if Morningstar provides the manager’s BA year, I subtract 21 from the year to get the birth year. For the performance of the fund, I use the Pastor, Stambaugh, and Taylor (2016) method. I subtract the corresponding monthly index return from the fund’s monthly gross return. Then I aggregate across months to obtain the annual outperformance. This method allows me to avoid estimating the beta, and some studies show that this method is better at controlling for risks than is the standard linear factor model. For the investment category of the fund, I use the Morningstar category. To measure the mutual fund managers’ connections to corporate I used two measures: (1) the number of connected board members and (2) the number of connected firms. The connection definition is CONNECTED(narrow). The dependent variable $Termination_{i,t}$ is set to one if the manager responsible for fund $i$ in January of year $t$ is no longer in charge of the fund at the beginning of year $t + 1$. I also construct a promotion dummy that equals 1 if the manager managed more funds in year $t$ than in year $t - 1$ and a demotion dummy that equals 1 if the
manager managed fewer funds in year $t$ than in year $t - 1$. The specification is as follows:

$$Y_{i,t} = \beta_0 + \beta_1 Connection + \beta_2 Alpha_{i,t} + \beta_3 Alpha_{i,t} \times (MgrAge_{i,t} - \bar{Age})$$

$$+ \beta_4 Alpha_{i,t-1} + \beta_5 Alpha_{i,t-2} + \beta_6 MgrAge_{i,t} + InvCat_j + \mu_t + \epsilon_{i,t}, \quad (A.1)$$

where $Y$ is a dummy for termination, promotion, or demotion. $InvCat_j$ is the fund investment category fixed effects and $\mu_t$ is the time fixed effects. Since the dependent variable is a binary outcome, I use a probit model to estimate it. Table A.14 presents the results. For the three types of career outcomes, connections to the firm are not statistically significant. This means that connections with board members are not important enough to affect the career path of mutual fund managers.

[INSERT TABLE A.14 HERE]

**A1.2. Connected Mutual Fund and Say on Pay Votes**

The Dodd-Frank Act mandates that, starting from 2011, the management of public firms should submit proposals about the top five executives’ compensation and shareholders can cast advisory vote of “yes” or “no” on the pay of the company’s top named executives during the prior fiscal year. Table A.15 reports the results. Connected mutual funds vote in favor of management.

[INSERT TABLE A.15 HERE]

**A1.3. Resistance to Hedge Fund Activism**

Activists go after companies with vulnerabilities. Most companies have some vulnerabilities, which could be related to financial performance, such as missing quarterly numbers, a stagnant stock price, or comparatively weak revenue growth. Weaknesses, such as missing quarterly numbers, could be a result of a firm engaging in long-term plans. But hedge funds
could step in and “increase” the efficiency of the firm. The activists’ agenda usually includes increasing leverage, returning excess cash to shareholders, selling off non-core corporate assets, cutting operating costs, and demanding representation on the board. Some of those procedures could be harmful to the firm’s long-term value such as innovation.

In this section, I check connected mutual fund’s trading behavior when facing hedge fund activism. I find that during the period when hedge funds accumulate their shares, connected funds stick with their equity in the firm, whereas non-connected funds tend to sell.

I obtain hedge fund activism data from Professor Wei Jiang. Brav, Jiang, Partnoy, and Thomas (2008) provide details on how they construct the dataset. The current hedge fund activism data I am using cover the period from 1994 to 2011. Following the literature, I use the date that the hedge fund files a Schedule 13D as the event date. I aggregate data into quarterly frequency. For a given firm, in a given quarter, if there is at least one hedge fund targeting them, I label this quarter as an event quarter for the firm.\footnote{In the data, there exists the case in which more than one hedge fund activism event occurred at a firm within a quarter. There are only 63 cases.} In the final sample, there are 1,822 hedge fund activism events affecting 1,452 different firms.

Table A.16 shows the results. In column (1), the dependent variable is the change in $\text{ConHold}$ from quarter $t - 3$ to quarter $t$, and, in column (2), it is the change in $\text{Non} - \text{ConHold}$ from quarter $t - 3$ to quarter $t$. $\text{Non} - \text{ConHold}$ is the difference between institutional holdings and $\text{ConHold}$. The dependent variable in columns (3) to (5) is the difference between the above two variables. All the dependent variables are in basis points. Hedge fund activism is a dummy variable that equals 1 if the firm in quarter $t$ has a hedge fund activism. Control variables include firm size, book-to-market ratio, and the previous year’s total stock return. Industry refers to the Fama-French 12-industry classifications. Standard errors are clustered at the firm level.

In column (1), I find that connected holdings are not affected by hedge fund activism. This means that when the hedge fund tries to accumulate shares in the firm, connected mutual
fund are less likely to sell. Meanwhile, for the non-connected funds, their shares significantly
decrease in the period before hedge fund activism actually happens. The difference between
the two types of funds trading can amount to 1% of a firm’s equity. Usually, hedge funds
accumulate around 5% of a firm’s equity before they initiate their activism. This means
that institutional investors who are not connected to the firm provide at least 1%. The
results are robust to include more control variables and industry and time fixed effects.

[INSERT TABLE A.16 HERE]

A1.4. Passive Funds in the Data

Table A.17 provides information about active and passive funds in my dataset. Passive
funds compose small amount of AUM in the data.

A1.5. Possion Model and Control Function Method

The basic assumption is that $y$’s distribution, conditional on all $x$, follows the Poisson
distribution. This is similar to the MLE world. We assume that, conditional on all the $x$, $y$
follows a normal distribution. Then for the Poisson distribution, it is as follows:

$$f(y|x) = \exp[-\mu(x)][\mu(x)]^y/y!,$$

where $\mu(x)$ decides the mean of the Poisson distribution. We can choose a parametric form
for the $\mu(x) = \exp(x\beta)$. This is because the $\mu(x)$ has to be positive all the time.

In my situation, because I assume $ConHold$ is endogenous. I will use a control function
method to deal with the endogeneity problem in the exponential regression setting.

First, assume the structural model follows:

$$E(y|z, ConHold, c_1) = \exp(z_1\delta_1 + ConHold\gamma_1 + c_1), \quad (A.2)$$
where $c_1$ is the unobserved latent variable that affects innovation and potentially causes the endogeneity problem. $z_1$ is a $1 \times L_1$ subset of $z$ containing unity. Here, we assume $z$ contains both an exogenous regressor, $z_1$ (i.e., all the control variables), and instrument variables, $z_2$ (i.e., ConMFM). On the one hand, $z$ is assumed to be uncorrelated with $c_1$. On the other hand, I allow the correlation between $ConHold$ and $c_1$. Further, I assume $ConHold$ has a linear reduced form satisfying the following assumption:

$$ConHold = z\Pi_2 + v_2.$$  \hfill (A.3)

This is exactly the first-stage regression in the IV analysis. In addition, I assume that $(c_1, v_2)$ is independent of $z$ and

$$c_1 = \rho_1 v_2 + \epsilon_1,$$  \hfill (A.4)

where $\epsilon_1$ is independent of $v_2$. Note that $ConHold$ is exogenous if and only if $\rho_1 = 0$.

Then I can write the original structural model as

$$E(y|z, ConHold, c_1) = exp(z_1\delta_1 + ConHold\gamma_1 + v_2\rho_1),$$  \hfill (A.5)

and estimating this equation using the standard quasi-maximum likelihood estimate (QMLE) method could allow me to obtain consistent estimates for $\delta_1, \gamma_1, \rho_1$. In implementation, estimate expression (A.3) and then compute the residual $\hat{v}_2 = ConHold - z\hat{\Pi}_2$. Then plug the residual $\hat{v}$ along with $z_1$ and $ConHold$ into equation (A.5).
A1.6. Citation Adjustment Method

I use the patents filed between 1980 and 1990 as my sample for estimation. For each technology class, I first compute the total number of citations. Then I aggregate the number of citations a specific technology class patents receives in each subsequent year after being filed. I divided each year’s citation number by the total number of citations to obtain the adjustment factors, which are presented in Table A.11. Then I adjust the number of citations that a patent currently has based on the number of years the patent has been filed and the adjustment factors.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Measures of innovation</strong></td>
<td></td>
</tr>
<tr>
<td><em>PatCite</em></td>
<td>Total number of citations received by the filed patents for firm <em>i</em> in year <em>t</em></td>
</tr>
<tr>
<td><em>PatApp</em></td>
<td>Total number of patents filed (and eventually granted) by firm <em>i</em> in year <em>t</em></td>
</tr>
<tr>
<td><em>KPSS</em></td>
<td>Total dollar value of innovation produced by firm <em>i</em> in year <em>t</em>, based on stock market’s reaction to the patents granted news, scaled by firm’s total asset</td>
</tr>
<tr>
<td><strong>Measures of connected holdings and other control variables</strong></td>
<td></td>
</tr>
<tr>
<td><em>ConHold</em></td>
<td>Fraction of company <em>i</em>’s equity held by connected mutual funds in year <em>t</em></td>
</tr>
<tr>
<td>Inst holding</td>
<td>Calculated as the mean of the four quarterly institutional holdings reported on form 13F</td>
</tr>
<tr>
<td>DED inst holding</td>
<td>The institutional holdings by dedicated institutional investors. Classification refers to Bushee (2001)</td>
</tr>
<tr>
<td>QIX inst holding</td>
<td>The institutional holdings by quasi-indexer institutional investors. Classification refers to Bushee (2001)</td>
</tr>
<tr>
<td>TRA inst holding</td>
<td>The institutional holdings by transitory institutional investors. Classification refers to Bushee (2001)</td>
</tr>
<tr>
<td>ROA</td>
<td>Return-on-assets ratio defined as operating income before depreciation (oibdp) divided by total assets (at)</td>
</tr>
<tr>
<td>Vol</td>
<td>Firm <em>i</em>’s idiosyncratic volatility measured as the standard deviation for the idiosyncratic part of firm return defined as ( \tilde{r} = r - r_m ), where <em>r</em> is firm’s daily return, ( r_m ) is the daily market return</td>
</tr>
<tr>
<td>R&amp;D stock</td>
<td>Following Hall, Jaffe, and Trajtenberg (2005), this variable is calculated using a perpetual inventory method: ( G_t = R_t + (1 - \delta)G_{t-1} ), where ( R_t ) is the R&amp;D expenditure in year <em>t</em> and ( \delta = 0.15 ), the private depreciation rate of knowledge</td>
</tr>
<tr>
<td>PPE/Assets</td>
<td>Property, plant, &amp; and equipment (ppent) divided by total assets (at)</td>
</tr>
<tr>
<td>Leverage</td>
<td>Book value of debt (dlc+dltt) divided by total assets (at)</td>
</tr>
<tr>
<td>Q</td>
<td>Tobin’s Q, calculated as market value of equity ((csho*prcc_c) plus total assets (at) minus book value of equity (ceq) minus balance sheet deferred taxes (pstkl, set to 0 if missing), divided by total assets (at)</td>
</tr>
<tr>
<td>R&amp;D/Asset</td>
<td>Research and development expenditures ((xrd+rdip) divided by total assets (at), set to 0 if missing</td>
</tr>
<tr>
<td>Capex/Assets</td>
<td>Capital expenditure ((capx) scaled by total assets (at)</td>
</tr>
<tr>
<td>Age</td>
<td>The number of years since initial public offering (IPO) (proxied by the first appearance in Compustat)</td>
</tr>
<tr>
<td>Sale</td>
<td>Firm’s sale</td>
</tr>
</tbody>
</table>
Table A.2: Days between Patent Filing Date and Grant Date

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>P1</th>
<th>P5</th>
<th>P25</th>
<th>P50</th>
<th>P75</th>
<th>P95</th>
<th>P99</th>
<th>N</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990–1999</td>
<td>714</td>
<td>385</td>
<td>235</td>
<td>316</td>
<td>495</td>
<td>652</td>
<td>854</td>
<td>1,289</td>
<td>1,925</td>
<td>352,370</td>
</tr>
<tr>
<td>2000–2010</td>
<td>1,141</td>
<td>561</td>
<td>292</td>
<td>433</td>
<td>732</td>
<td>1,037</td>
<td>1,453</td>
<td>2,171</td>
<td>2,828</td>
<td>730,670</td>
</tr>
</tbody>
</table>

This table reports the days between patent filing date and patent grant date in the dataset for the period from 1980 to 2010. Here, I only include patents granted to public firms. (Other type of entities, for example, universities, also have patents, but are not included). Means, standard deviations, and key percentiles are provided in the table.

Table A.3: Number of Connections

<table>
<thead>
<tr>
<th>Broad connection: Same university</th>
<th>Narrow connection: Same univ/degree/time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harvard University</td>
<td>Harvard University</td>
</tr>
<tr>
<td>University of Pennsylvania</td>
<td>University of Pennsylvania</td>
</tr>
<tr>
<td>Columbia University</td>
<td>University of Chicago</td>
</tr>
<tr>
<td>University of Chicago</td>
<td>Stanford University</td>
</tr>
<tr>
<td>New York University</td>
<td>University of California, Berkeley</td>
</tr>
<tr>
<td>Stanford University</td>
<td>Columbia University</td>
</tr>
<tr>
<td>University of California, Berkeley</td>
<td>University of Wisconsin, Madison</td>
</tr>
<tr>
<td>Northwestern University</td>
<td>Princeton University</td>
</tr>
<tr>
<td>University of Michigan</td>
<td>University of Michigan</td>
</tr>
<tr>
<td>University of Wisconsin, Madison</td>
<td>Yale University</td>
</tr>
<tr>
<td>Cornell University</td>
<td>University of Virginia</td>
</tr>
<tr>
<td>Massachusetts Institute of Technology</td>
<td>University of Illinois</td>
</tr>
<tr>
<td>University of Texas at Austin</td>
<td>Cornell University</td>
</tr>
<tr>
<td>Yale University</td>
<td>New York University</td>
</tr>
<tr>
<td>University of Virginia</td>
<td>Dartmouth College</td>
</tr>
</tbody>
</table>

This table reports the number of connections for the top 15 educational institutions according to two definitions: broad connection broad, that is, mutual fund managers and corporate board members attend the same school, and narrow connection, that is, mutual fund managers and corporate board members attended the same school for the same type of degree at the same time.
Table A.4: Top 10 Most Connected University

<table>
<thead>
<tr>
<th>Connected firms</th>
<th>Connected funds</th>
</tr>
</thead>
<tbody>
<tr>
<td>Harvard University</td>
<td>780</td>
</tr>
<tr>
<td>University of Pennsylvania</td>
<td>344</td>
</tr>
<tr>
<td>Stanford University</td>
<td>330</td>
</tr>
<tr>
<td>Columbia University</td>
<td>257</td>
</tr>
<tr>
<td>Yale University</td>
<td>237</td>
</tr>
<tr>
<td>New York University</td>
<td>226</td>
</tr>
<tr>
<td>Massachusetts Institute of Technology</td>
<td>196</td>
</tr>
<tr>
<td>Princeton University</td>
<td>194</td>
</tr>
<tr>
<td>University of Michigan</td>
<td>182</td>
</tr>
<tr>
<td>University of Chicago</td>
<td>180</td>
</tr>
</tbody>
</table>

This table reports the time-series average of connected firms and connected funds for the sample from 1980 to 2003. *Connected* means board members and fund managers obtained degrees from the same university.

Table A.5: Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>P5</th>
<th>P25</th>
<th>P50</th>
<th>P75</th>
<th>P95</th>
</tr>
</thead>
<tbody>
<tr>
<td>PatCite</td>
<td>74.373</td>
<td>321.398</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>320.396</td>
</tr>
<tr>
<td>PatApp</td>
<td>4.078</td>
<td>18.000</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>16</td>
</tr>
<tr>
<td>KPSS</td>
<td>0.027</td>
<td>0.122</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.135</td>
</tr>
<tr>
<td>ConHold\textsuperscript{broad} (%)</td>
<td>0.401</td>
<td>1.743</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2.509</td>
</tr>
<tr>
<td>ConHold\textsuperscript{narrow} (%)</td>
<td>0.020</td>
<td>0.251</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Inst holding(%)</td>
<td>18.556</td>
<td>23.822</td>
<td>0</td>
<td>0</td>
<td>6.238</td>
<td>32.549</td>
<td>69.019</td>
</tr>
<tr>
<td>DED inst holding(%)</td>
<td>2.663</td>
<td>5.624</td>
<td>0</td>
<td>0</td>
<td>0.000</td>
<td>2.905</td>
<td>13.671</td>
</tr>
<tr>
<td>QIX inst holding(%)</td>
<td>11.304</td>
<td>15.281</td>
<td>0</td>
<td>0</td>
<td>2.972</td>
<td>19.227</td>
<td>44.016</td>
</tr>
<tr>
<td>TRA inst holding(%)</td>
<td>4.119</td>
<td>7.355</td>
<td>0</td>
<td>0</td>
<td>0.092</td>
<td>5.511</td>
<td>19.840</td>
</tr>
<tr>
<td>ROA</td>
<td>0.049</td>
<td>23.593</td>
<td>-0.425</td>
<td>0.015</td>
<td>0.108</td>
<td>0.171</td>
<td>0.278</td>
</tr>
<tr>
<td>Vol</td>
<td>0.608</td>
<td>1.964</td>
<td>0.044</td>
<td>0.124</td>
<td>0.279</td>
<td>0.630</td>
<td>2.037</td>
</tr>
<tr>
<td>R&amp;D stock</td>
<td>44.448</td>
<td>619.595</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>47.201</td>
</tr>
<tr>
<td>PPE/Assets</td>
<td>0.306</td>
<td>0.231</td>
<td>0.032</td>
<td>0.121</td>
<td>0.248</td>
<td>0.440</td>
<td>0.782</td>
</tr>
<tr>
<td>Leverage</td>
<td>0.240</td>
<td>0.213</td>
<td>0</td>
<td>0.049</td>
<td>0.207</td>
<td>0.369</td>
<td>0.646</td>
</tr>
<tr>
<td>Q</td>
<td>1.990</td>
<td>1.813</td>
<td>0.725</td>
<td>1.019</td>
<td>1.365</td>
<td>2.147</td>
<td>5.582</td>
</tr>
<tr>
<td>R&amp;D/Asset</td>
<td>0.011</td>
<td>0.046</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.070</td>
</tr>
<tr>
<td>CAPX/Asset</td>
<td>0.074</td>
<td>0.078</td>
<td>0.004</td>
<td>0.024</td>
<td>0.049</td>
<td>0.093</td>
<td>0.237</td>
</tr>
<tr>
<td>Observations</td>
<td>113,503</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Panel B: Samples Sorted by Connected Holdings

<table>
<thead>
<tr>
<th>Variable</th>
<th>$\text{ConHold}^{\text{narrow}} &gt; 0$</th>
<th>$\text{ConHold}^{\text{narrow}} = 0$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>SD</td>
</tr>
<tr>
<td>$\text{PatCite}$</td>
<td>403.563</td>
<td>758.187</td>
</tr>
<tr>
<td>$\text{PatApp}$</td>
<td>22.344</td>
<td>42.523</td>
</tr>
<tr>
<td>$\text{KPSS}$</td>
<td>0.132</td>
<td>0.293</td>
</tr>
<tr>
<td>$\text{ConHold}^{\text{broad}}$(%)</td>
<td>5.320</td>
<td>4.672</td>
</tr>
<tr>
<td>$\text{ConHold}^{\text{narrow}}$(%)</td>
<td>0.566</td>
<td>1.208</td>
</tr>
<tr>
<td>Inst holding(%)</td>
<td>60.895</td>
<td>19.023</td>
</tr>
<tr>
<td>DED inst holding(%)</td>
<td>8.475</td>
<td>7.019</td>
</tr>
<tr>
<td>QIX inst holding(%)</td>
<td>37.327</td>
<td>13.062</td>
</tr>
<tr>
<td>TRA inst holding(%)</td>
<td>14.689</td>
<td>9.798</td>
</tr>
<tr>
<td>ROA</td>
<td>0.142</td>
<td>0.124</td>
</tr>
<tr>
<td>Vol KPSS</td>
<td>0.265</td>
<td>0.306</td>
</tr>
<tr>
<td>R&amp;D stock</td>
<td>508.795</td>
<td>2070.772</td>
</tr>
<tr>
<td>PPE/Assets</td>
<td>0.287</td>
<td>0.205</td>
</tr>
<tr>
<td>Lev</td>
<td>0.223</td>
<td>0.185</td>
</tr>
<tr>
<td>Q</td>
<td>2.553</td>
<td>1.952</td>
</tr>
<tr>
<td>R&amp;D/Asset</td>
<td>0.035</td>
<td>0.062</td>
</tr>
<tr>
<td>CAPX/Asset</td>
<td>0.065</td>
<td>0.057</td>
</tr>
<tr>
<td>Observations</td>
<td>4,059</td>
<td>109,444</td>
</tr>
</tbody>
</table>

Panel A reports the summary statistics for the full sample of variables constructed based on the sample of U.S. public firms from 1980 to 2003. Panel B separately shows statistics for those with positive CONNECTED(narrow) holdings (left) and those with no CONNECTED(narrow) holdings (right). Table A.1 provides variable definitions.

Table A.5: Age distribution

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>SD</th>
<th>P10</th>
<th>P25</th>
<th>P50</th>
<th>P75</th>
<th>P90</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fund manager</td>
<td>47.04</td>
<td>10.16</td>
<td>35</td>
<td>40</td>
<td>45</td>
<td>54</td>
<td>62</td>
</tr>
<tr>
<td>Board member</td>
<td>54.16</td>
<td>9.90</td>
<td>42</td>
<td>47</td>
<td>54</td>
<td>61</td>
<td>67</td>
</tr>
</tbody>
</table>

This table displays the distribution of fund manager and board member age in my dataset as of year 2003.
Table A.6: Connected Holdings and Innovation (Poisson Model)

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ConHold</td>
<td>362.32***</td>
<td>465.44***</td>
</tr>
<tr>
<td></td>
<td>(74.50)</td>
<td>(79.76)</td>
</tr>
<tr>
<td>Year FEs</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Firm FEs</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Control variables</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Observations</td>
<td>113,503</td>
<td>113,503</td>
</tr>
</tbody>
</table>

Table A.1 lists variable definitions. Control variables include annual institutional holdings, log sale, log firm age, R&D/Asset, PPE/Asset, leverage, Q, ROA, and CAPX/Asset. The estimated model is the Poisson model with firm and year fixed effects, and robust standard errors are calculated and provided in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table A.7: Connected Holdings and Innovation Controlling for CEO FE

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ConHold</td>
<td>37.41***</td>
<td>66.91**</td>
<td>9.13***</td>
</tr>
<tr>
<td></td>
<td>(14.49)</td>
<td>(27.84)</td>
<td>(2.86)</td>
</tr>
<tr>
<td>Year FE</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Firm FE</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>University FE</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>CEO FE</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Control variables</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.89</td>
<td>0.83</td>
<td>0.68</td>
</tr>
<tr>
<td>Observations</td>
<td>29,219</td>
<td>29,219</td>
<td>29,219</td>
</tr>
</tbody>
</table>

This table reports the results of regressions of the innovation outcomes on connected holdings and other control variables. The difference between this table and Table 1.2 is that in this table, I also control for CEO fixed effects. Table A.1 provides variable definitions. Control variables include annual institutional holdings, log sale, log firm age, R&D/Asset, PPE/Asset, leverage, Q, ROA, and CAPX/Asset. University FE are defined in Section 1.3.2.2. The estimated model is two-stage least-squares model with firm and year fixed effects. Robust standard errors were calculated and are provided in parentheses. ***, **, and * indicate that the coefficient is significant at the 1%, 5%, and 10% level, respectively.
Table A.8: Connected Holdings (Management Only) and Innovation

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>LnPatApp</th>
<th>LnPatCite</th>
<th>LnKPSS</th>
<th>LnPatApp</th>
<th>LnPatCite</th>
<th>LnKPSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ConHold</td>
<td>50.32**</td>
<td>82.50**</td>
<td>10.88***</td>
<td>42.26***</td>
<td>59.47***</td>
<td>8.83***</td>
</tr>
<tr>
<td>Year FEs</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Firm FEs</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>University FEs</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Control variables</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Observations</td>
<td>110,663</td>
<td>110,663</td>
<td>110,663</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

This table reports the results of regressions of innovation outcome on connected holdings and other control variables for two types of connection definitions. Table A.1 lists the variable definitions. Control variables include annual institutional holdings, log sale, log firm age, R&D/Asset, PPE/Asset, leverage, Q, ROA, and CAPX/Asset. University FEs are defined in Section 1.3.2.2. The estimated model is a two-stage least-squares model with firm and year fixed effects, and robust standard errors are calculated and provided in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table A.9: Connected Holdings (Broad) and Innovation

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>LnPatApp</th>
<th>LnPatCite</th>
<th>LnKPSS</th>
</tr>
</thead>
<tbody>
<tr>
<td>ConHoldBroad</td>
<td>5.45***</td>
<td>4.95**</td>
<td>0.53***</td>
</tr>
<tr>
<td>Year FE</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Firm FE</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>University FE</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>CEO FE</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Control variables</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.85</td>
<td>0.77</td>
<td>0.63</td>
</tr>
<tr>
<td>Observations</td>
<td>110,663</td>
<td>110,663</td>
<td>110,663</td>
</tr>
</tbody>
</table>

This table reports the results of regressions of the innovation outcomes on connected holdings and other control variables. The difference between this table and Table 1.2 is that in this table, I use broad connections to compute connected holdings. Broad connection means board members and mutual fund managers went to the same school. Table A.1 provides variable definitions. Control variables include annual institutional holdings, log sale, log firm age, R&D/Asset, PPE/Asset, leverage, Q, ROA, and CAPX/Asset. University FE are defined in Section 1.3.2.2. The estimated model is two-stage least-squares model with firm and year fixed effects. Robust standard errors were calculated and are provided in parentheses. ***, **, and * indicate that the coefficient is significant at the 1%, 5%, and 10% level, respectively.
Table A.10: Connected Holdings versus Institutional Holdings

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>lnPatApp</td>
<td>42.26***</td>
<td>59.47***</td>
<td>8.83***</td>
<td>lnKPSS</td>
<td></td>
<td></td>
</tr>
<tr>
<td>lnPatCite</td>
<td>(12.83)</td>
<td>(22.60)</td>
<td>(2.10)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lnKPSS</td>
<td>0.041</td>
<td>0.133**</td>
<td>-0.018***</td>
<td>0.079***</td>
<td>0.201***</td>
<td>-0.008***</td>
</tr>
<tr>
<td>lnPatApp LnPatCite</td>
<td>(0.001)</td>
<td>(0.062)</td>
<td>(0.004)</td>
<td>(0.029)</td>
<td>(0.0058)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>lnKPSS LnKPSS</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Year FEs</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Firm FEs</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>University FEs</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Control variables</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>R-squared</td>
<td>0.84</td>
<td>0.77</td>
<td>0.59</td>
<td>0.82</td>
<td>0.74</td>
<td>0.58</td>
</tr>
<tr>
<td>Observations</td>
<td>110,663</td>
<td>110,663</td>
<td>110,663</td>
<td>110,663</td>
<td>110,663</td>
<td>110,663</td>
</tr>
</tbody>
</table>

Table A.11: Citation adjustment factors

<table>
<thead>
<tr>
<th>Age</th>
<th>Tech class 1</th>
<th>Tech class 2</th>
<th>Tech class 3</th>
<th>Tech class 4</th>
<th>Tech class 5</th>
<th>Tech class 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0083</td>
<td>0.0059</td>
<td>0.0047</td>
<td>0.0071</td>
<td>0.0061</td>
<td>0.0058</td>
</tr>
<tr>
<td>1</td>
<td>0.0442</td>
<td>0.0303</td>
<td>0.0259</td>
<td>0.0384</td>
<td>0.0350</td>
<td>0.0324</td>
</tr>
<tr>
<td>2</td>
<td>0.1049</td>
<td>0.0735</td>
<td>0.0609</td>
<td>0.0931</td>
<td>0.0871</td>
<td>0.0798</td>
</tr>
<tr>
<td>3</td>
<td>0.1725</td>
<td>0.1252</td>
<td>0.1051</td>
<td>0.1559</td>
<td>0.1480</td>
<td>0.1345</td>
</tr>
<tr>
<td>4</td>
<td>0.2403</td>
<td>0.1832</td>
<td>0.1540</td>
<td>0.2201</td>
<td>0.2097</td>
<td>0.1911</td>
</tr>
<tr>
<td>5</td>
<td>0.3083</td>
<td>0.2450</td>
<td>0.2067</td>
<td>0.2834</td>
<td>0.2724</td>
<td>0.2492</td>
</tr>
<tr>
<td>6</td>
<td>0.3733</td>
<td>0.3092</td>
<td>0.2627</td>
<td>0.3447</td>
<td>0.3326</td>
<td>0.3067</td>
</tr>
<tr>
<td>7</td>
<td>0.4365</td>
<td>0.3749</td>
<td>0.3212</td>
<td>0.4035</td>
<td>0.3912</td>
<td>0.3642</td>
</tr>
<tr>
<td>8</td>
<td>0.4954</td>
<td>0.4394</td>
<td>0.3810</td>
<td>0.4596</td>
<td>0.4465</td>
<td>0.4199</td>
</tr>
<tr>
<td>9</td>
<td>0.5521</td>
<td>0.5012</td>
<td>0.4414</td>
<td>0.5127</td>
<td>0.5006</td>
<td>0.4740</td>
</tr>
<tr>
<td>10</td>
<td>0.6072</td>
<td>0.5611</td>
<td>0.5033</td>
<td>0.5639</td>
<td>0.5548</td>
<td>0.5285</td>
</tr>
<tr>
<td>11</td>
<td>0.6589</td>
<td>0.6191</td>
<td>0.5641</td>
<td>0.6151</td>
<td>0.6076</td>
<td>0.5835</td>
</tr>
<tr>
<td>12</td>
<td>0.7094</td>
<td>0.6747</td>
<td>0.6241</td>
<td>0.6645</td>
<td>0.6592</td>
<td>0.6380</td>
</tr>
<tr>
<td>13</td>
<td>0.7577</td>
<td>0.7278</td>
<td>0.6873</td>
<td>0.7144</td>
<td>0.7114</td>
<td>0.6908</td>
</tr>
<tr>
<td>14</td>
<td>0.8042</td>
<td>0.7801</td>
<td>0.7482</td>
<td>0.7633</td>
<td>0.7612</td>
<td>0.7439</td>
</tr>
<tr>
<td>15</td>
<td>0.8479</td>
<td>0.8297</td>
<td>0.8042</td>
<td>0.8110</td>
<td>0.8108</td>
<td>0.7964</td>
</tr>
<tr>
<td>16</td>
<td>0.8864</td>
<td>0.8728</td>
<td>0.8553</td>
<td>0.8560</td>
<td>0.8571</td>
<td>0.8455</td>
</tr>
<tr>
<td>17</td>
<td>0.9213</td>
<td>0.9117</td>
<td>0.8988</td>
<td>0.8986</td>
<td>0.8984</td>
<td>0.8922</td>
</tr>
<tr>
<td>18</td>
<td>0.9517</td>
<td>0.9455</td>
<td>0.9364</td>
<td>0.9370</td>
<td>0.9355</td>
<td>0.9319</td>
</tr>
<tr>
<td>20</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
<td>1.0000</td>
</tr>
</tbody>
</table>

This table provides the citation adjustment factors. Technology classes are defined in Hall, Jaffe, and Trajtenberg (2001). More specifically, technology class: 1, Chemical (excluding Drugs); 2, Computers and Communications (C&C); 3, Drugs and Medical (D&M); 4, Electrical and Electronics (E&E); 5, Mechanical; and 6, Others. See appendix 1 in Hall, Jaffe, and Trajtenberg (2001) for more details. Age is the number of years the patent has been filed. I make the assumption that a patent accumulates all its citations after 20 years.
Table A.12: Daily Trading Volume and Connected Holdings

<table>
<thead>
<tr>
<th></th>
<th>Daily Trading Volume</th>
<th>$ConHold^{narrow}$</th>
<th>$ConHold^{broad}$</th>
<th>Inst Holding</th>
<th>Observations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>P25</td>
<td>P75</td>
<td>Mean</td>
<td>Mean</td>
</tr>
<tr>
<td>1990</td>
<td>0.0050</td>
<td>0.0026</td>
<td>0.0061</td>
<td>0.0041</td>
<td>0.0147</td>
</tr>
<tr>
<td>2000</td>
<td>0.0099</td>
<td>0.0053</td>
<td>0.0116</td>
<td>0.0054</td>
<td>0.0579</td>
</tr>
<tr>
<td>2003</td>
<td>0.0097</td>
<td>0.0112</td>
<td>0.0055</td>
<td>0.0064</td>
<td>0.0672</td>
</tr>
</tbody>
</table>

This table reports daily trading volume for stocks with positive $ConHold^{narrow}$. Daily trading volume is defined as the daily stock sold divided by the total shares outstanding. For each stock in each year, I compute the annual average. I also report the 25th and 75th percentile of the daily trading volume for a given stock in a given year. Table A.1 provides all variable definitions.

Table A.13: Summary Statistics for Termination and Connectedness

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>SD</th>
<th>Median</th>
</tr>
</thead>
<tbody>
<tr>
<td>Termination</td>
<td>0.105</td>
<td>0.307</td>
<td>0</td>
</tr>
<tr>
<td>Promotion</td>
<td>0.069</td>
<td>0.255</td>
<td>0</td>
</tr>
<tr>
<td>Demotion</td>
<td>0.036</td>
<td>0.186</td>
<td>0</td>
</tr>
<tr>
<td>No. connected firm</td>
<td>22.329</td>
<td>28.707</td>
<td>12</td>
</tr>
<tr>
<td>No. connected director</td>
<td>18.307</td>
<td>21.920</td>
<td>10</td>
</tr>
<tr>
<td>No. fund managed</td>
<td>1.576</td>
<td>1.049</td>
<td>1</td>
</tr>
<tr>
<td>Performance</td>
<td>0.019</td>
<td>0.115</td>
<td>0.006</td>
</tr>
<tr>
<td>Mutual fund manager age</td>
<td>45.633</td>
<td>8.874</td>
<td>45</td>
</tr>
<tr>
<td>Observations</td>
<td>3.416</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The data are at the fund-year frequency. I only include single manager-managed funds. Termination is a dummy equaling 1 when either the fund manager stops managing the fund or when the fund exits. Promotion is a dummy equaling 1 if a manager manages more funds in the next year than in this year. Demotion is a dummy equaling 1 if a manager manages fewer funds in the next year than in this year. According to the definition of $CONNECTED^{(narrow)}$, the number of connected firms is the number of connected firms to funds in that year. A similar definition for the number of connected directors. There are cases in which one manager could manage more than one fund, so I count the number of funds managed. Alpha is calculated as the difference between a fund’s gross return and its corresponding index return.
Table A.14: Relationship between Termination and Connectedness

<table>
<thead>
<tr>
<th>Dependent variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. connected firms</td>
<td>-0.000120</td>
<td>-0.000235</td>
<td>-0.00210</td>
<td>-0.000516</td>
<td>-0.000359</td>
<td>-0.00247</td>
</tr>
<tr>
<td></td>
<td>(0.00125)</td>
<td>(0.00136)</td>
<td>(0.00191)</td>
<td>(0.00164)</td>
<td>(0.00178)</td>
<td>(0.00247)</td>
</tr>
<tr>
<td>No. connected directors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alpha&lt;sub&gt;t&lt;/sub&gt;</td>
<td>-0.585</td>
<td>0.376</td>
<td>-0.0934</td>
<td>-0.583</td>
<td>0.376</td>
<td>-0.0929</td>
</tr>
<tr>
<td></td>
<td>(0.384)</td>
<td>(0.416)</td>
<td>(0.581)</td>
<td>(0.384)</td>
<td>(0.416)</td>
<td>(0.581)</td>
</tr>
<tr>
<td>Alpha&lt;sub&gt;t&lt;/sub&gt; * Age</td>
<td>0.0357</td>
<td>-0.0547</td>
<td>0.0687</td>
<td>0.0361</td>
<td>-0.0546</td>
<td>0.0679</td>
</tr>
<tr>
<td></td>
<td>(0.0353)</td>
<td>(0.0390)</td>
<td>(0.0499)</td>
<td>(0.0353)</td>
<td>(0.0389)</td>
<td>(0.0500)</td>
</tr>
<tr>
<td>Alpha&lt;sub&gt;t−1&lt;/sub&gt;</td>
<td>-0.477</td>
<td>0.529</td>
<td>-0.870</td>
<td>-0.478</td>
<td>0.528</td>
<td>-0.875</td>
</tr>
<tr>
<td></td>
<td>(0.375)</td>
<td>(0.390)</td>
<td>(0.579)</td>
<td>(0.375)</td>
<td>(0.390)</td>
<td>(0.578)</td>
</tr>
<tr>
<td>Alpha&lt;sub&gt;t−2&lt;/sub&gt;</td>
<td>-0.987**</td>
<td>0.667*</td>
<td>-0.210</td>
<td>-0.986**</td>
<td>0.667*</td>
<td>-0.215</td>
</tr>
<tr>
<td></td>
<td>(0.387)</td>
<td>(0.403)</td>
<td>(0.542)</td>
<td>(0.387)</td>
<td>(0.403)</td>
<td>(0.542)</td>
</tr>
<tr>
<td>Age</td>
<td>-0.0110**</td>
<td>0.00631</td>
<td>0.00123</td>
<td>-0.0110**</td>
<td>0.00628</td>
<td>0.00870</td>
</tr>
<tr>
<td></td>
<td>(0.00447)</td>
<td>(0.00487)</td>
<td>(0.00616)</td>
<td>(0.00445)</td>
<td>(0.00486)</td>
<td>(0.00614)</td>
</tr>
<tr>
<td>Year FEs</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Investment category FEs</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>

This table reports the probit model estimations of equation A.1. Standard errors are in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level, respectively.
Table A.15: Connected Mutual Fund Votes on “Say-on-Pay” Issue

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent variable</td>
<td>0.0366</td>
<td>0.0386</td>
<td>0.1085**</td>
<td>0.1083**</td>
</tr>
<tr>
<td>if vote with board</td>
<td>(0.0338)</td>
<td>(0.0342)</td>
<td>(0.0435)</td>
<td>(0.0438)</td>
</tr>
<tr>
<td>recommendation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Connected(broad)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control variables</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Industry FEs</td>
<td>no</td>
<td>yes</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Pseudo R-squared</td>
<td>0.0000</td>
<td>0.0175</td>
<td>0.2885</td>
<td>0.2918</td>
</tr>
<tr>
<td>Observations</td>
<td>249,205</td>
<td>249,205</td>
<td>249,205</td>
<td>249,205</td>
</tr>
<tr>
<td>Marginal effect on</td>
<td>0.0061</td>
<td>0.0063</td>
<td>0.0124</td>
<td>0.0122</td>
</tr>
<tr>
<td>Connected(broad)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In this table, the sample is restricted to “say-on-pay” proposals. Each observation represents the vote of a mutual fund on a proposal made at a company’s shareholder meeting. The dependent variable equals 1 if the mutual fund’s vote follows the board’s recommendation. Connected(broad) is a dummy variable that equals 1 if the mutual fund manager went to the same university as at least one of the firm’s board members. Control variables include ISS recommendation, firm size, book-to-market ratio, the previous year’s total stock return, fund size, fund turn ratio, fund expense ratio, and fund family size. Industry refers to the Fama-French 12-industry classifications. Standard errors clustered at firm level are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.
Table A.16: Mutual Fund Trading and Hedge Fund Activism

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ConHold</td>
<td>Non − ConHold</td>
<td>Diff</td>
<td>Diff</td>
<td>Diff</td>
</tr>
<tr>
<td>Hedge fund activism</td>
<td>-0.3539</td>
<td>-105.8***</td>
<td>105.49***</td>
<td>79.94***</td>
<td>99.69***</td>
</tr>
<tr>
<td>(0.9923)</td>
<td>(24.10)</td>
<td>(24.13)</td>
<td>(27.46)</td>
<td>(31.17)</td>
<td></td>
</tr>
<tr>
<td>Controls</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Fixed effects</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Observations</td>
<td>274,510</td>
<td>274,510</td>
<td>274,510</td>
<td>164,755</td>
<td>164,755</td>
</tr>
</tbody>
</table>

The dependent variable in column (1) is the change in ConHold from quarter $t − 3$ to quarter $t$, and, in column (2), it is the change in Non − ConHold from quarter $t − 3$ to quarter $t$. Non − ConHold is the difference between institutional holdings and ConHold. The dependent variable in columns (3) to (5) is the difference between the above two variables. All the dependent variables are in basis points. Hedge fund activism is a dummy variable that equals 1 if the firm in quarter $t$ has a hedge fund activism. Control variables include firm size, book-to-market ratio, and the previous year’s total stock return. Industry refers to the Fama-French 12-industry classifications. Standard errors clustered at the firm level are in parentheses. ***, **, and * indicate statistical significance at the 1%, 5%, and 10% level, respectively.

Table A.17: Active Funds versus Passive Funds

<table>
<thead>
<tr>
<th></th>
<th>Number of funds</th>
<th>Total AUM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1) Active</td>
<td>Passive</td>
</tr>
<tr>
<td>1980</td>
<td>279</td>
<td>1</td>
</tr>
<tr>
<td>1985</td>
<td>440</td>
<td>3</td>
</tr>
<tr>
<td>1990</td>
<td>801</td>
<td>14</td>
</tr>
<tr>
<td>1995</td>
<td>1,449</td>
<td>37</td>
</tr>
<tr>
<td>2000</td>
<td>2,589</td>
<td>76</td>
</tr>
<tr>
<td>2005</td>
<td>3,204</td>
<td>95</td>
</tr>
<tr>
<td>2010</td>
<td>3,031</td>
<td>93</td>
</tr>
</tbody>
</table>

This table shows the number of funds and total AUM (in billions) for active funds and passive funds in my sample. Passive funds are defined as funds that contain “Index” or “index” in their fund name.
This figure displays the total number of patent applications (that are eventually granted) in the dataset by years. A censor problem shows up from year 2002 or 2003 until the end of the sample.
This figure displays the annual growth rate of the total number of patent applications (defined as the current year’s total number of patents divided by last year’s total number of patents) in the dataset from 1980 to 2010.
The graph displays the total AUM of mutual funds that exclusively hold domestic stocks as a fraction of the total AUM of all matched mutual funds. The data series covers quarter 1 of 1980 to quarter 4 of 2003.
This histogram plots the distance in miles between the headquarters of mutual fund companies and the headquarters of public firms. The vertical red line indicates 100 miles. According to Coval and Moskowitz (2001), if the distance between the headquarters of mutual fund companies and that of public firms is smaller than 100 km (which is less than 100 miles), the holdings will be defined as local holdings.
A2. Appendix for Chapter 2

A2.1. Investor beliefs

We use the Kalman filter to derive investor belief about manager skill. Let \( y_{j,t} \equiv r_{j,t} + D(s_{j,t}; \eta) \). By (2.1), we have

\[
y_{j,t} = a_{j,t} + \varepsilon_{j,t}.
\]

We can treat this as the measurement equation in a state space representation. The state equation is a simple AR(1) process for \( a_{j,t} \) as specified in (2.2). Obtaining Equation (2.3) and (2.4) is simply a matter of applying the Kalman filter. In particular, \( \tilde{a}_{j,t} \) is the one period ahead prediction of the state, and \( \tilde{\sigma}_{j,t} \) is the variance of that prediction.

A2.2. Optimality of cut-off strategy

Here, we provide a few details on how to derive the optimal search strategy for the investors. Fix an investor in a period. For notational simplicity, we suppress the subscript \( i \) and subscript \( t \). The Bellman equation for the dynamic problem is

\[
V(u^*) = \max \left\{ u^*, \quad -c + \int_{-\infty}^{+\infty} V(\max\{u^*, u\}) d\Psi(u) \right\}.
\]

Consider a cutoff strategy that stops at any \( u > \bar{u} \). With such a strategy, \( V(u^*) = u^* \) for all \( u^* > \bar{u} \). On the other hand, the value for \( u^* \leq \bar{u} \) should be given by

\[
V(u^*) = \sum_{t=0}^{+\infty} \Psi(\bar{u})^t \left[ 1 - \Psi(\bar{u}) \right] \left[ \int_{(\bar{u}, \infty)} ud\Psi(u) \right] (t+1)c - \int_{(\bar{u}, \infty)} ud\Psi(u) - c \left[ 1 - \Psi(\bar{u}) \right] \sum_{t=0}^{+\infty} \Psi(\bar{u})^t (t+1) = \frac{1}{1 - \Psi(\bar{u})} \int_{(\bar{u}, \infty)} ud\Psi(u) - c \left[ 1 + 2\Psi(\bar{u}) + 3\Psi(\bar{u})^2 + 4\Psi(\bar{u})^3 + \ldots \right] = \frac{1}{1 - \Psi(\bar{u})} \left[ \int_{(\bar{u}, \infty)} ud\Psi(u) - c \right].
\]

(A.6)
On the right side of the first line, $\Psi(\overline{u})[1 - \Psi(\overline{u})]$ is the probability that the investor does not stop for $t$ periods and then stops. Multiplying this probability is the expectation of the sampled $u$ that triggers the stop minus the incurred search costs of $t + 1$ periods.

Most importantly, notice that (A.6) is a constant that does not depend on $u^*$. In addition, we must have $V(\overline{u}) = \overline{u}$. Equating (A.6) with $\overline{u}$ gives us the expression for $\overline{u}$ that we gave in the main text:

$$c = \int_{(\overline{u},\infty)} (u - \overline{u})d\Psi(u).$$

With $\overline{u}$ thus defined, the value function can be written as

$$V(u^*) = \max\{u^*, \overline{u}\}.$$

We can verify that it satisfies the Bellman equation, as for $u^* \leq \overline{u}$,

$$-c + \int_{-\infty}^{+\infty} V(\max\{u^*, u\}) d\Psi(u) = -c + \int_{-\infty}^{+\infty} \max\{u, \overline{u}\}d\Psi(u)$$

$$= -c + \overline{u} + \int_{(\overline{u},\infty)} (u - \overline{u})d\Psi(u)$$

$$= \overline{u},$$

and for $u^* > \overline{u}$,

$$-c + \int_{-\infty}^{+\infty} V(\max\{u^*, u\}) d\Psi(u) = -c + \int_{-\infty}^{+\infty} \max\{u, u^*\}d\Psi(u)$$

$$= -c + u^* + \int_{(u^*,\infty)} (u - u^*)d\Psi(u)$$

$$< u^*.$$
A2.3. Market shares

To facilitate subsequent derivations, here we define a fund-specific cutoff \( f_j, j = 0, 1, ..., N, \) where
\[
f_j = \sum_{k=0}^{N} \psi_k (u_k - u_j) \cdot 1\{u_k > u_j\}.
\]

Notice that \( u_j = \bar{u}(f_j) \). So, if \( c_i > f_j \), then \( u_j > \bar{u}(c_i) \). In other words, if an investor’s search cost is larger \( f_j \), he will stop searching once he finds fund \( j \). With these funds’ specific cutoffs, we can derive closed-form expressions for market shares: first, for the fund with the lowest utility, then, for the fund with the second lowest utility, etc. Let \( \tau \) be a permutation on \( \{0, 1, ..., N\} \) such that \( u_{\tau(0)} \leq u_{\tau(1)} \leq ... \leq u_{\tau(N)} \). As a result, \( f_{\tau(0)} \geq f_{\tau(1)} \geq ... \geq f_{\tau(N)} \).

Any investor who has a search cost that is higher than \( f_{\tau(0)} \) will not make a second search beyond the free search. Then, among all of these investors, with \( \psi_{\tau(0)} \) probability, they will find fund \( \tau(0) \) (the “worst” fund). Nevertheless, they will invest in fund \( \tau(0) \). No one else will invest with fund \( \tau(0) \). So the market share for fund \( \tau(0) \) is
\[
s_{\tau(0)} = \psi_{\tau(0)} \left[ 1 - G(f_{\tau(0)}) \right],
\]
where \( G \) is the c.d.f. for the distribution of \( c_i \) in the population.

Two kinds of investors will buy fund \( \tau(1) \). The first kind is the investors with \( c_i > f_{\tau(0)} \) that find fund \( \tau(1) \) in the free search. They have no choice but to invest. The second kind is investors with \( f_{\tau(0)} \geq c_i > f_{\tau(1)} \). For these investors to invest in fund \( \tau(1) \), they could have found it in the free search, or have found \( \tau(0) \) in the free search and \( \tau(1) \) in the second search, or have found \( \tau(0) \) in the first two searches and \( \tau(1) \) in the third search, and so forth. The total probability of these events is
\[
\psi_{\tau(1)} + \psi_{\tau(0)} \psi_{\tau(1)} + \psi_{\tau(0)}^2 \psi_{\tau(1)} + ... = \frac{\psi_{\tau(1)}}{1 - \psi_{\tau(0)}}.
\]
So the market share for fund $\tau(1)$ is

$$s_{\tau(1)} = \psi_{\tau(1)} \left[ 1 - G(f_{\tau(0)}) \right] + \frac{\psi_{\tau(1)}}{1 - \psi_{\tau(0)}} \left[ G(f_{\tau(0)}) - G(f_{\tau(1)}) \right]$$

$$= \psi_{\tau(1)} \left[ 1 + \frac{\psi_{\tau(0)} G(f_{\tau(0)}) - G(f_{\tau(1)})}{1 - \psi_{\tau(0)}} \right].$$

We can follow this line of deduction to obtain closed-form expressions for the market shares of all funds. For $j \geq 2$,

$$s_{\tau(j)} = \psi_{\tau(j)} \left[ 1 + \sum_{k=0}^{j-1} \frac{\psi_{\tau(k)} G(f_{\tau(k)})}{(1 - \psi_{\tau(0)} - \cdots - \psi_{\tau(k-1)})(1 - \psi_{\tau(0)} - \cdots - \psi_{\tau(k)})} \right].$$

### A2.4. Investor welfare

The previous section provides the proof that the optimal search strategy is a cutoff strategy. In this section we compute investor $i$’s welfare for a given search cost $c_i$. First, we denote $\bar{u}(c_i)$ as the reservation level of utility for the investor $i$. Investor $i$ will only accept the funds which provide utilities higher or equal to $\bar{u}(c_i)$, so the expected utility for the potentially accepted funds is

$$\int_{\bar{u}(c_i)}^{+\infty} u d\Psi(u) \frac{1}{1 - \Psi(\bar{u}(c_i))}.$$

As to the search cost, the probability that the investor will conduct $t$ searches beyond the
free search is \((1 - \Phi)^t\), so the expected total search cost is

\[
c[1 - \Psi(\bar{u})] \sum_{t=0}^{+\infty} \Psi(\bar{u})^t = c [1 - \Psi(\bar{u})] \{ [\Psi(\bar{u}) + \Psi(\bar{u})^3 + ...] + \}
\]

\[
\{ [\Psi(\bar{u})^2 + \Psi(\bar{u})^3 + ...] + ... \} = c [1 - \Psi(\bar{u})] \{ \frac{1}{1 - \Psi(\bar{u})} + \frac{\Psi(\bar{u})}{1 - \Psi(\bar{u})} + ... \}
\]

\[
= c \frac{\Psi(\bar{u})}{1 - \Psi(\bar{u})}
\]

where \(\bar{u}\) is \(\bar{u}(c_j)\). Combining the two parts together, we have the expression for investor’s expected welfare.

### A2.5. Frictionless case

Here we derive the limiting case of our model when the search costs go to zero, \(\lambda \to 0\). We fix a time period \(t\) and suppress the subscript \(t\) throughout the derivation.

First, notice that the active funds must provide the same utility, \(u_j = u'\) for some \(u'\) for all \(j \in \{1, \ldots, N\}\). To see this, suppose that some \(j\) has a utility that is strictly smaller than another fund. Because the investors do not incur search costs, no one will buy \(j\). This means \(s_j \to 0\), which under the log specification of the decreasing returns to scale implies that \(u_j \to +\infty\), a contradiction. By the same argument, one can show that \(u' \geq -p_0\).

Let us first look at the case that \(u' > -p_0\) for all \(j \in \{1, \ldots, N\}\). The outside good will have zero market share. So

\[
\sum_{j=1}^{N} s_j = 1.
\]

In addition, from the utility specification (2.5), we have

\[
s_j = e^{\eta_j - \eta(p_j + u')}.
\]

Putting the two above equations together, we can find the solution for \(u'\) and plug it back
into the last equation to obtain:

\[ s_j = \frac{e^{\frac{1}{\eta} \tilde{a}_j - \frac{1}{\eta \gamma} p_j}}{\sum_{k=1}^{N} e^{\frac{1}{\eta} \tilde{a}_k - \frac{1}{\eta \gamma} p_k}}. \tag{A.7} \]

Next, consider the case where \( u' = -p_0 \). The size of an active fund will be at the point where the decreasing returns to scale drives its utility to be the same as the index fund: this is the key idea of Berk and Green (2004). From the utility specification (2.5), we have

\[ s_j = e^{\frac{1}{\eta} \tilde{a}_j - \frac{1}{\eta \gamma} (p_j - p_0)}. \tag{A.8} \]

For this case, we must have \( \sum_{j=1}^{N} s_j \leq 1 \), which translates into

\[ -p_0 \geq \eta \gamma \log \left( \sum_{k=1}^{N} e^{\frac{1}{\eta} \tilde{a}_k - \frac{1}{\eta \gamma} p_k} \right). \tag{A.9} \]

In other words, if this condition on the prices holds, then the market shares are given by (A.8), otherwise the market shares are given by (A.7).

We can derive the pricing behavior of funds given these market share equations. Each fund chooses \( p_j \) to maximize \( s_j(p_j - b_j - mc_j) \). Suppose that condition (A.9) holds so that \( s_j \) is given by (A.8), then the first order condition implies a uniform markup of \( \eta \gamma \) across the active funds, or

\[ p_j = \eta \gamma + b_j + mc_j. \]

If these prices satisfy condition (A.9), then we have a Nash-Bertrand equilibrium in which the index fund has a positive market share.

A2.6. Uniqueness of the fixed point

In this section, we show that the fixed point defined as

\[ F_t [p_t, b_t, \tilde{a}_t - \eta \log (M_t s_t), x_t, \xi_t, p_{0,t}; \Theta] = s_t \]
is unique. For notational simplicity, we suppress the subscript \(t\). We use a result from Kennan (2001), which provides the uniqueness of a fixed point under R-concavity and the quasi-increasing condition. We first need to show that
\[
F[p, b, \tilde{a} - \eta \log (Ms), x, \xi, p_0; \Theta] - s
\]
as a function of \(s\) is strictly R-concave, i.e., for any \(0 < z < 1\), we have
\[
F[p, b, \tilde{a} - \eta \log (Ms), x, \xi, p_0; \Theta] > zs. \tag{A.10}
\]
Notice that \(\tilde{a} - \eta \log (zM) = \tilde{a} - \eta \log (Ms) + \eta \log (z^{-1})\), which increases the utility for all the active funds by the same amount \(\eta \log (z^{-1})\). This is equivalent to lowering the utility of the outside good (i.e., index fund) by \(\eta \log (z^{-1})\). So in the following, we show that lowering the utility of the outside good increases the market share of every active fund. This will imply (A.10).

Recall that we have for \(j = 0, 1, 2, ..., N\), the market share for \(\tau(j)\) equals the summation of \(j + 1\) terms:
\[
F_{\tau(j)} = \psi_{\tau(j)}[1 - G(f_{\tau(0)})] + \frac{\psi_{\tau(j)}}{1 - \psi_{\tau(0)}}[G(f_{\tau(0)}) - G(f_{\tau(1)})] + 
\]
\[
... + \frac{\psi_{\tau(j)}}{1 - \psi_{\tau(0)} - ... - \psi_{\tau(j-1)}}[G(f_{\tau(j-1)}) - G(f_{\tau(j)})].
\]
where
\[
f_j = \sum_{k=0}^{N} \psi_k (u_k - u_j) \cdot 1\{u_k > u_j\}.
\]
Suppose that there is a small incremental on \(u_0\). Formally, let \(u'_0 = u_0 + \Delta, u'_j = u_j\) for all \(j \neq 0\). Consider the case where \(u_0\) is not equal to any \(u_j, j \neq 0\). Then we can take \(\Delta\) small enough such that the ranking of \(\{u_j\}_{j=0}^{N}\) and the ranking of \(\{u'_j\}_{j=0}^{N}\) are identical, which means that the same permutation \(\tau\) can be used. Let \(k\) be such that \(\tau(k) = 0\), i.e., the
index fund is ranked at the \( k \)th position. We have

\[
f'_j = \begin{cases} 
  f_j + \psi_0 \Delta, & \text{if } u_j < u_0; \\
  f_j, & \text{if } u_j > u_0; \\
  f_0 - (1 - \psi_0(0) - \cdots - \psi_\tau(k)) \Delta & \text{if } j = 0.
\end{cases}
\]

Then, for a general \( j \neq k \), \( F'_\tau(j) \) is the summation of \( j + 1 \) terms:

\[
F'_\tau(j) = \psi_\tau(j) \left[ 1 - G(f_\tau(0) + \psi_0 \Delta) \right] + \frac{\psi_\tau(j)}{1 - \psi_\tau(0)} \left[ G(f_\tau(0) + \psi_0 \Delta) - G(f_\tau(1) + \psi_0 \Delta) \right] + \cdots + \frac{\psi_\tau(j)}{1 - \psi_\tau(0) - \cdots - \psi_\tau(k-1)} \left[ G(f_\tau(k-1) + \psi_0 \Delta) - G(f_\tau(k) - (1 - \psi_\tau(0) - \cdots - \psi_\tau(k)) \Delta) \right] + \frac{\psi_\tau(j)}{1 - \psi_\tau(0) - \cdots - \psi_\tau(j-1)} \left[ G(f_\tau(j-1)) - G(f_\tau(j)) \right].
\]

Hence,

\[
\lim_{\Delta \to 0} \frac{F'_\tau(j) - F_\tau(j)}{\Delta} = -\psi_\tau(j) \psi_0 G'(f_\tau(0)) + \frac{\psi_\tau(j) \psi_0}{1 - \psi_\tau(0)} \left[ G'(f_\tau(0)) - G'(f_\tau(1)) \right] + \cdots + \frac{\psi_\tau(j)}{1 - \psi_\tau(0) - \cdots - \psi_\tau(k-1)} \left[ \psi_0 G'(f_\tau(k-1)) + G'(f_\tau(k)) \cdot (1 - \psi_\tau(0) - \cdots - \psi_\tau(k)) \right] + \frac{\psi_\tau(j)}{1 - \psi_\tau(0) - \cdots - \psi_\tau(k)} G'(f_\tau(k)) \cdot (1 - \psi_\tau(0) - \cdots - \psi_\tau(k)).
\]

Combining the last two terms, we have

\[
\lim_{\Delta \to 0} \frac{F'_\tau(j) - F_\tau(j)}{\Delta} = -\psi_\tau(j) \psi_0 G'(f_\tau(0)) + \frac{\psi_\tau(j) \psi_0}{1 - \psi_\tau(0)} \left[ G'(f_\tau(0)) - G'(f_\tau(1)) \right] + \cdots + \frac{\psi_\tau(j)}{1 - \psi_\tau(0) - \cdots - \psi_\tau(k-1)} \left[ G'(f_\tau(k-1)) - G'(f_\tau(k)) \right].
\]

Under the exponential specification of \( G \), we know that (i) \( G' > 0 \); (ii) \( G'(f_\tau(k-1)) - G'(f_\tau(k))i0 \). With these two facts, it is easy to see that \( \lim_{\Delta \to 0} \frac{F'_\tau(j) - F_\tau(j)}{\Delta} < 0 \). So we have
essentially shown that when \( u_0 \) does not equal the utility of any other fund,

\[
\frac{\partial F_j}{\partial u_0} < 0, \ \forall j = 1, ..., N.
\]

Because there are only finite points at which \( u_0 \) becomes equal to the utility of some other fund, the above result implies that \( F_j \) is strictly decreasing in \( u_0 \) for all \( j \neq 0 \). In words, lowering the utility of the outside good increases the market share of every active fund, which is what we started out to show.

The second condition that we need to show in order to apply the result in Kennan (2001) is that \( F [p, b, \tilde{a} - \eta \log (Ms), x, \xi, p_0; \Theta] \), as a function of \( s \), is strictly radially quasiconcave. That is, for any \( s \) and \( s' \) where \( s_j = s'_j \) but \( s'_k \geq s_k \) for all \( k \neq j \), we have

\[
F_j [p, b, \tilde{a} - \eta \log (Ms'), x, \xi, p_0; \Theta] \geq F_j [p, b, \tilde{a} - \eta \log (Ms), x, \xi, p_0; \Theta].
\]

In other words, we need to show that when the utilities of all but one active fund decrease, the market share of this one active fund increases. To prove this, we only need to apply a similar argument as above to show that

\[
\frac{\partial F_j}{\partial u_k} < 0, \ \forall j, k = 1, ..., N \text{ and } j \neq k.
\]

except for possibly a finite set of points.

Lastly, by Theorem 1 from Kennan (2001) we show that if a positive fixed point exists, it is unique.
A2.7. Computation and estimation

Let $s_{j,t}$ be the observed share for fund $j$ in period $t$. Given a set of parameters, we can find the $\xi_t$ by matching our model predicted shares with the observed shares:

$$H_{j,t}(p_t, b_t, \tilde{a}_t, x_t, \xi_t, p_{0,t}; \Theta) = s_{j,t}, \quad (A.11)$$

where $\tilde{a}_t$ is obtained from (2.3) using the parameter values estimated in Section 2.4.1. Solving for $\xi_t$ can be done in a similar fashion as the contraction mapping in Berry et al. (1995). However, because $H_{j,t}$ requires fixed-point iteration to evaluate, this is computationally costly. Instead, solving for $\xi_t$ from

$$F_{j,t}[p_t, b_t, \tilde{a}_t, x_t, \xi_t, p_{0,t}; \Theta] = s_{j,t} \quad (A.12)$$

is generally faster, because $F_{j,t}$ has closed-form expressions as derived in Section 2.2.3. This amounts to plugging the observed $s_t$ in to the left hand side and searching for the value $\xi_t$ that makes $F_{j,t}$ equal to the observed $s_{j,t}$ for each $j$. Given the definition of $H_{j,t}$ in (2.10), solving (A.11) and solving (A.12) are equivalent.

A2.8. Components of sampling probability

There is a way to see the impact of various components of our model on capital misallocation. We redraw Figure 2.1 but remove each of the components (one at a time): $\xi$, fund age, fund family size, and marketing expenses. The solid black line is the BG model-implied “efficient” fund size. The blue line is the data. The dashed line plots the new fund size as predicted by the “restricted” model (in the sense of eliminating a particular heterogeneity but keeping the current equilibrium, i.e., without re-solving for all the funds’ best response strategies).

In the first figure, all the portfolios shift upwards in parallel. This is due to the noise introduced into fund size, which raises the log size on average (by the Jensen’s inequality).

In the second figure, all the portfolios’ average sizes shift downwards because fund age is
useful for informing investors. In the third figure, we can see the counterfactual is similar to the data, this means that fund family size is not very important in affecting fund size. The last figure plots the counterfactual fund size when there is no marketing. Interestingly, we see that the dashed line becomes steeper than observed in the data and thus closer to the “efficient” allocation. This means that marketing helps less skilled funds attain larger market share than they would otherwise.

[Insert Figure A.10 from Online Appendix Here]

A2.9. Standard errors

The standard errors can be computed by parametric bootstrap. The only element that we have to take as exogenous in the simulation is the existence of the funds over time (we do not have a model of entry and exit). The shocks that we need to generate include $\nu_{j,t}$, $\varepsilon_{j,t}$, $\xi_{j,t}$, $\zeta_{j,t}$, and $\omega_{j,t}$. The latter two shocks are highly correlated (as explained in Section 2.4.2,) and each shows persistence over time. One way to incorporate these is by using a VAR process. We can start at year $t = 1$, first take the $\tilde{a}_{j,1}$ as the prior beliefs, then compute the equilibrium prices, marketing expenses, and market shares, given the prior beliefs and a set of randomly drawn $\xi_{j,1}$'s. After this, we can move on to $t = 2$, first compute the belief $\tilde{a}_{j,2}$ based on the simulated $r_{j,1}$ and $s_{j,1}$, then compute the equilibrium given these beliefs and a set of $\xi_{j,2}$'s, and so on until the last period $T$. This provides us with a panel of simulated data on which we can apply our estimation algorithm. We run Monte Carlo experiments to verify that our estimator is able to recover the “true” parameters.

A2.10. Search model parameter sensitivity

In order to verify that our model estimation is well-specified, we report sensitivities of our parameter estimates to two key moments using a local measure developed in Andrews, Gentzkow and Shapiro (2017). This measure helps us assess how much the parameters change if moment conditions are violated. Many of the results in this paper rely on correctly estimating the demand effects of price and marketing. In our estimation, the demand
elasticities of price and marketing are identified from two behavioral assumptions (equation (2.14) and (2.15)). These two moment conditions require that on average mutual funds are setting their expense ratios and marketing expenditures given these price and marketing elasticities. The latter of the optimality conditions, in particular, might be violated on average if either the upper or the lower limit on the marketing expenditure is binding. Here we show how the parameter estimates will change if there are small systematic deviations from these optimal behaviors.

The results are provided in table A.21. The numbers can be interpreted as the percentage bias of the parameter estimates if a moment is violated by 1% of the standard deviation of $\omega$ or $\zeta$. For example, if the average $\zeta$ is not equal to zero but rather to 1% of the estimated standard deviation of $\zeta$, then the estimate for $\lambda$ would be downward biased by 0.65% from its true value, approximately. Overall, we find that the mean search cost $\lambda$ and the weight of gross outperformance in the utility function $\gamma$ are somewhat sensitive to the violations of the pricing moment condition (with sensitivities of $-0.64$ and $-0.54$, respectively, when using all funds in the sample). The other parameters are insensitive to the violations of this moment condition. The effect of marketing expenditure on sampling probability $\theta$ is also somewhat sensitive to the violation of marketing moment condition (with sensitivity of 0.32), while other parameters are not.

The latter sensitivity can be used to assess the degree to which bounds on marketing expenditures effect our parameter estimates. When funds are constrained in their ability to market by the SEC-imposed cap on 12b-1 fees (so that $\sum \omega_{j,t} > 0$), but the econometrician (mistakenly) assumes the moment condition holds with equality, the demand effects of marketing will be exaggerated, which manifests in the model as a smaller coefficient in front of the marketing expenses. However, this effect is not very large, as a 1% of the standard deviation increase in the moment condition translates into roughly one third of a percentage point increase in the marketing coefficient. More importantly, the effect is essentially the same for funds that are not at the upper bound (and similar for those at...
the lower bound of zero), suggesting that the impact of the binding constraints on the estimation of search model parameters is small.

[Insert Table A.21 Here]
A2.11. Data Appendix

In this appendix, we describe our dataset construction procedure. The raw data come from CRSP Survivor-Bias-Free US mutual fund dataset and Morningstar.

A2.12. Matching between CRSP and Morningstar

Our goal is to merge CRSP mutual fund dataset with Morningstar dataset. The identifiers that are common across these two datasets are ticker and CUSIP. However, in CRSP, the unique identifier is crsp\_fundno and in Morningstar, it is secid. Both identify a unique share class not a fund (for example, C share class of X fund). In this section, we create the one-to-one mapping between crsp\_fundno and ticker; crsp\_fundno and CUSIP; secid and ticker; secid and CUSIP. We follow Berk and van Binsbergen’s (hereafter BB) procedure as close as possible.

A2.12.1. crsp\_fundno and ticker mapping

We download the annual fund summary dataset from CRSP through Wharton Research Data Service (WRDS). The data spans from Jan 1961 to Dec 2015. There are 505,073 observations.

1. Out of 505,073 observations, there are 400 observations with same \{crsp\_fundno, year\} as other observations. These duplications happen due to multiple reports in the same year. Out of the 400 observations, there are 200 distinct crsp\_fundnos. We keep the observation with non missing expense ratio information and delete the others. Now we have 504,808 observations. After this step, we don’t have any observations with identical crsp\_fundno and year.

2. Out of 504,808 observations left, we have 86,793 obs for which ticker is missing. We follow BB’s steps to fill those. First, we identify all the unique pairs of \{crsp\_fundno, ticker\}. Here we first delete the observations with missing tickers. Then, we delete
the observations with duplicated pairs of \{crsp\_fundno, ticker\}. We got 53,278 unique pairs. We find that there are 5,425 pairs of which have the same crsp\_fundno but more than one ticker. We follow BB’s procedure: we keep the latest ticker which is the ticker with the most recent year. Then, we back fill all the previous ticker with that tickers. This step gives us 2,595 additional unique pairs of \{crsp\_fundno, ticker\}. Then, we add back the pairs from the non duplicated case, we have 50,448 unique \{crsp\_fundno, ticker\} pairs.

3. Up to this point, for each crsp\_fundno, there is only one ticker. But for each ticker there could be multiple crsp\_fundnos. Now we identify the tickers that have multiple crsp\_fundnos. There are actually 4,343 of them. We follow BB to treat those pairs as missing. Now we get 42,436 unique pairs of \{crsp\_fundno, ticker\}.

Feature: our dataset have one-to-one mapping between crsp\_fundno and ticker.

A2.12.2. crsp\_fundno and CUSIP mapping

According to Pstor, Stambaugh and Taylor 2015 (hereafter PST), CUSIP can match a lot of Morningstar funds to CRSP funds in addition to tickers. So we also clean the CUSIP in CRSP. In general, we conduct the exact same procedures as we did with the ticker. So in the following, we only report some key statistics.

1. Out of 505,073 observations, there are 120,837 observations with missing CUSIPs. After we do the back fill, the number of observations with missing CUSIP is reduced to 29,436.

2. Next, we identify the CUSIPs that have been used by multiple crsp\_fundno. There are 494 such CUSIPs. We set them to missing.

3. Lastly, we have 53,297 unique pairs of \{crsp\_fundno, CUSIP\}.

Feature: our dataset has one-to-one mapping between crsp\_fundno and CUSIP.
We merge the above two datasets together and keep all observations. Now a fund at least has a ticker or a CUSIP. It could have both. This leaves us with 54,911 unique crsp.fundnos.

A2.12.3. Morningstar data

We start from the fund_ops file that we have gotten from Morningstar. This dataset contains the Morningstar Category, Fund Family name and other information. It has 55,571 obs.

We only keep the domestic well-diversified equity mutual funds. We follow the method provided in PST data appendix in identifying this type of funds.

1. We first identify the observations with duplicated fund_name and delete them. We also delete the funds with no Morningstar category which corresponds to additional 661 funds.

2. Then, we identify the bond fund, international fund, sector fund, target date fund, real estate fund, other non-equity fund. The definition and method are provided in PST. Now we are left with 23,592 funds.

3. We delete the funds with neither a ticker nor a CUSIP. We are left with 21,580 funds.

A2.12.4. Merge between CRSP and Morningstar

Our goal is to get one-to-one mapping between crsp.fundno and secid through ticker or CUSIP. For details on secid, please check PST.

We use the CRSP dataset that has unique pairs between crsp.fundno and ticker or CUSIP to merge with Morningstar. First, we merge on ticker. We get 12,412 matches.

A small issue in the Morningstar dataset is that the CUSIP is 9 digit while in CRSP, the CUSIP is 8 digit. Following the instruction on WRDS, we get rid of the last digit of CUSIP in Morningstar dataset.

Then, we merge on CUSIP. We get 17,488 matches.
Finally, we take the union of the two types of matches and we delete the duplicated pairs of \{crsp\_fundno, secid\}. We have 17,658 unique crsp\_fundno and secid pairs.

**A2.13. CRSP dataset clean**

We merge the above identified 17,658 unique observations with annual CRSP Fund Summary dataset and keep the merged observations. We denote this dataset as the *baseline* dataset. It has 169,488 observations at fund share class/year level.

**A2.14. Correct TNA**

As pointed out in PST, before 1993, a lot of the funds in CRSP dataset report their assets under management (AUM or TNA, same meaning) at a quarterly or annual frequency. Meanwhile, most of the funds report their returns at monthly frequency. When we aggregate variables such as returns and expense ratios across share classes, we need the monthly TNA information. Starting from the raw dataset of mutual funds monthly returns downloaded from WRDS, we merge it with the 17,658 unique crsp\_fundno and keep the merged observations. This gives us 2,149,498 observations (covers year from 1962 to 2016). Then we do the following correction:

1. If there is no TNA information for a given fund for any month in a year, we delete this year.

2. For the funds who report TNA at an annual frequency (with only one non-missing TNA per year), we replace the other 11 month’s TNAs with the non-missing TNA.

3. For the funds who report their TNA at quarterly frequency (with only one non-missing TNA per quarter), we replace the missing values of TNA with the TNA in that quarter. For example, if a fund reports TNA at month 3 for quarter 1, we replace month 1 and month 2 TNA as month 3 TNA.

4. We delete the observations where TNA is zero, negative.
5. We delete the observations with missing TNA or monthly return.

6. We also delete duplicated observations for the same crsp_fundno in the same month.

After this correction, we have 2,018,242 observations with non missing monthly return and TNA. (In the data appendix, we use TNA and AUM interchangeably.)

A2.14.1. Inflation adjustment

To make the TNAs comparable across time, we take inflation into consideration. We download the Consumer Price Index from FRED, Federal Reserve Economic Data provided by St. Louis Fed. The series we used is Consumer Price Index for All Urban Consumers: All Items\(^2\). This series is at the monthly frequency and it is seasonally adjusted. We convert it to annual frequency by keeping each year’s December’s value as this year’s value for CPI. Then, we use year 2015 as the baseline year to adjust all other year’s TNAs. For example, the CPI value in 1970 is 39.6 and the CPI value in 2015 is 238.3. Then, all the monthly TNAs in 1970 are multiplied by 6 (\(= 238.3/39.6\)) to make them comparable to TNAs in 2015.

A2.15. Vanguard Index Fund

As proposed in BB, index funds from Vanguard are the most accessible index funds to the average investors. In our paper, we use all the equity index funds from Vanguard, combined, as the outside good. Within baseline dataset, we first drop all the institutional share classes: drop the funds if inst_fund== Y, sharetype == ”Inst” or fund name contains ”Institutional Shares” or ”Institutional Class”.

In order to identify index funds, we use a simple two-step procedure following PST:

1. If either CRSP or Morningstar indicates the fund is an index fund, we label this fund

\(^2\)The url is https://fred.stlouisfed.org/series/CPIAUCSL
as index fund.

2. If a fund’s name contains words such as ‘Index’ or ‘index’, we label this fund as index fund.

Then, we check whether the fund name or the fund sponsor’s name contains "Vanguard". There are 552 such observations, i.e., 552 Vanguard index fund observations.

We fill the missing value of expense ratio using the fund’s life time average. Then, we delete the observations with missing expense ratio. This gives us 494 share class/year observations. We merge it with fund monthly dataset which gives us 6,279 share class/month observations.

In each month, we get the total assets under management (inflation adjusted, for details see section “Inflation adjustment”), asset weighted mean of management fees, returns, expense ratios, turnover ratios, and 12b-1 fees. Then, for each month we only keep one observation for the Vanguard index fund. Then, for each year we keep first month’s value as the the Vanguard fund’s annual variable. We have 40 observations from year 1976 to 2015.

A2.16. Active Funds Cleaning

From the baseline dataset, we drop all the institutional share classes.

Then, we drop all the index funds. For methods, please check section "Vanguard Index Fund". Further we only keep the funds with the following Morningstar categories: Large Blend, Large Growth, Large Value, Mid-Cap Blend, Mid-Cap Growth, Mid-Cap Value, Small Blend, Small Growth, Small Value and Aggressive Allocation. We call this dataset Active Fund dataset. It has 87,842 observations at share class/year level.

A2.16.1. Front Load

To construct “effective” 12b-1 fees, we need information about fund’s front load. We downloaded the front load dataset from CRSP. The total number of observations is 101,848.
1. In CRSP mutual fund front load dataset, for each crsp_fundno there is a pricing schedule for the front load (i.e., for certain amount of initial investment, the fund will charge certain percentage of front load. Generally, the more the investment, the lower the front load.). For each pricing schedule we only keep the maximum level of front load.

2. Then, we delete the observations with front load equals 0. That leaves us with 19,626 observations.

3. We delete observations with front load smaller than 0. There are 30 of them.

4. There are 288 cases that a fund has more than one change in front load in one year. We delete them.

5. We expand the front load dataset to a crsp_fundno year style. This step gives us 108,818 observations.

Next, we merge the above front load data set with the Active Fund dataset. We set missing front loads to zero.

A2.16.2. “Effective” 12b-1 fees and expense ratio adjustment

We combine fund’s 12b-1 fee and front load to create an item we call “effective” 12b-1 fee. For the fund share class with missing value of expense ratio or 12b-1, we use the time series mean of the expense ratio or 12b-1 to replace the missing value. We set the missing 12b-1 fee to 0 and replace the observation with 12b-1 fee larger than 1% to 1%.

For fund \( j \) in year \( t \), if a C share class exists, we replace all the other share classes’ expense ratios and 12b-1 fees with the C share class’s data. The C share class is the class that charges no front load fees but has higher expense ratios and 12b-1 fees. We replace other share classes’ expense ratios and 12b-1 fees with the C share class’s data on the assumption

\[ ^3 \text{Because in the raw dataset, each entry has start year and end year.} \]
\[ ^4 \text{The expense ratio or 12b-1 fee smaller than 0 or equals -99 was set to missing before the replacing.} \]
that mutual fund investors are indifferent towards different share classes. This assumption is valid if all investors have the same investment horizons. If no C share class exists in the fund, then for all the other share classes, we take the sum of the share class’s 12b-1 fees and the annualized front load for that share class and use it as the effective 12b-1 fees. For this case, we also increase the expense ratio by the amount of the annualized front load. Following Sirri and Tufano (1998), we annualize the front load by dividing it by 7, implicitly assuming that it is amortized over 7 years.

The way to identify whether some share classes belong to the same fund is by using MS_fundid from Morningstar. Also we make sure that the “effective” 12b-1 fee is not larger than 1%. We also did the same adjustment to expense ratio. Finally, we drop the observation with expense ratio greater than 5% and we drop the observations with expense ratio smaller than the sum of 5bps and “effective” 12b-1 fee. We only keep the observations later than 1964 (include 1964).

In Figure A.7 panel A, we plot the ratio of marketing expenses to total fees for active funds. We can see that this ratio is relatively stable from 1992 to 2015, at around 41%. Figure A.7 panel B plots the aggregate time series of the total amount of marketing expenses in billion dollars. The mean is around 8.5 billion dollars per annum. Marketing expenses are substantial both in absolute terms and relative to the industry’s total revenues.

A2.16.3. Construct monthly return dataset

Starting from Active Fund dataset, we merge the monthly return dataset into it. Here the monthly TNAs are already corrected in the ”Correct TNA” section. We also convert TNA into real terms using the procedures described in section “Inflation adjustment”.

The gross return for each share class is the sum of net return (mret) and one twelfth
of expense ratio. (Here the expense ratio we use is from raw data instead of the one we generated in the previous section. We replace the missing value of expense ratio by the time series mean. Expense ratio smaller or equal to zero are treated as missing. The reason why we use this unadjusted expense ratio is because we want our gross return to be comparable to other papers which study mutual fund performance in this fashion.)

We aggregate the gross return, net return, 12b-1 fee and expense ratio to fund level by using each share class’s TNA as weight. We use MS_fundid as the identifier for a fund (not a share class). For each fund, each month, we only keep one observation.

We also clean the turnover ratio and management fee using the same procedure (in this section) as we clean the expense ratio. Here we further impose that expense ratio is larger than 20 bps. PST uses 15 bps as the lower bound for the expense ratio for active funds. After this step we have 516,849 observations.

A2.16.4. General cleaning

Starting from the dataset in the previous step, we keep the the observations with TNA ≥15 million dollars (this is the threshold used in PST). Then, we keep only the cases where the fund in a given year has 12 observations. Next, we keep month 1’s observation (to convert the dataset from monthly to annual.). We check whether there is a gap in this annual dataset. For example for fund \( j \), suppose it has the annual data from 1996 to 2000 and 2002 to 2010. Then, we will delete this fund, i.e. all the 14 (=5+9) observations. We also drop the active funds from Vanguard since Vanguard is organized as a client-owned structure which might not be properly described by profit maximization which is what we have in the model. Now for all the active funds, we have 27,621 observations. It is at fund/year level. Now we append the Vanguard Index Fund data into the above dataset. There is additional 40 fund/year observations. This almost finishes all the data cleaning part. We call it almost there dataset.
A2.16.5. Performance adjustment

Starting with the *almost there* dataset, we merge the monthly return dataset from section “Construct Monthly Return Dataset”. We keep the merged observations which is 331,452.

To adjust for risk exposure, we use various versions of asset pricing models. The monthly factor returns are downloaded from Ken French’s website. Then, we use the following models to adjust the returns: CAPM, Fama-French 3 factor model, Fama-French-Carhart Model, Fama-French 5 factor model. The dependent variable is fund’s excess return which is the difference between gross monthly return of the fund and monthly risk free rate. To increase the accuracy of beta estimation, for each fund we use all the fund’s returns to estimate the betas. Then, we subtract the predicted return (betas multiplied by the factors’ returns) from the fund’s excess return. This gives us the monthly alphas.

Then, we aggregate alphas together by sum across 12 months to get annual alphas. And we merge the annual alphas to the *almost there* dataset to get our final dataset.

A2.16.6. Variable calculation

Total market is the sum of AUM for all the funds (active and passive) in a given year. Market share is ratio between fund’s AUM and Total market. Index fund price is the expense ratio of the Vanguard index fund. Family size is the number of funds in the same fund family. Fund family is identified by using Family variable from Morningstar. Fund age is the number of years since the fund first shows up in the dataset.
Panel A plots the total assets under management (AUM) of all the Vanguard equity index funds. The data series starts from 1975 when Vanguard launching its S&P 500 index fund. In year 2015, the total AUM reaches 600 billions of dollars. We find all Vanguard funds by fund names (contains 'Vanguard' or 'vanguard'). We manually screen out the non equity index funds by checking fund’s prospectus. The funds’ assets under management are inflated to 2015 dollars by using Consumer Price Index downloaded from FRED.

Panel B plots the asset weighted mean expense ratio for all the Vanguard equity index funds. The unit is in basis point. This series covers the period from 1975 to 2015. We can see a significant drop from over 60 bp in 1975 to under 10 bp in 2015.
Figure A.6: Histogram of Effective Marketing Expenses

The figure plots the histogram of effective marketing expenses for the main sample covering 1964 to 2015. We define the marketing expenses in the following way: for fund $j$ in year $t$, if a C share class exists, we replace all the other share classes expense ratios and 12b-1 fees with the C share class’s data. If no C share class exists in the fund, then for all the other share classes, we take the sum of the share class’s 12b-1 fee and the annualized front load for that share class and use it as the effective 12b-1 fee. For this case, we also increase the expense ratio by the amount of the annualized front load. Lastly, within a fund, across share classes, we aggregate the effective 12b-1 fee by the AUM of each share class to get the fund level effective 12b-1 fee. About 45.7% of the observations are binding at the upper bound, 1% level. And about 23.7% of the observations are binding at 0%.
Figure A.7: Total Marketing

Panel A: Total marketing expenses as a fraction of expense ratio for active funds

Panel B: Total marketing expenses for the active funds

Panel A plots the ratio between total marketing expenses and total expense ratios for active funds.
Panel B plots the total marketing expenses (in billions dollars) for all the active equity mutual funds in the U.S. from 1992 to 2015. The funds’ assets under management are inflated to 2015 dollars by using Consumer Price Index downloaded from FRED.
The figure presents the prior distribution of management skill (a). The vertical line marks the mean of expense ratio in our data, 1.66%. Approximately 71% of funds have management skills higher than the mean expense ratio. The parameter values are $\mu = 0.0305$ (mean of prior) and $\kappa = 0.0241$ (std of prior).

The figure plots the mean of log fund size (fund size is measured in million dollars) for portfolio of funds formed on net skill for the no-marketing equilibrium. The expense ratio is outcome of the counterfactual experiment. We compute fund size implied by the generalized Berk and Green (2004) model using the ratio between net skill and the degree of decreasing returns to scale (BG). The black line plots the mean of log fund size implied by BG. The blue line plots the mean of log fund size generated by our search model in the counterfactual equilibrium for each portfolio. The purple line plots the mean of log fund size generated by our search model in the counterfactual equilibrium with no foc errors for each portfolio.
The four figures plot the mean of log fund size (fund size is measured in million dollars) for portfolio of funds formed on net skill (defined as fund skill level $\tilde{a}$ minus expense ratio $p$). We compute the Berk and Green model implied fund size using the ratio between net skill and degree of decreasing returns to scale. The fund skills are estimated using our generalized Berk and Green model. For more details, please check section 2.2.1. The black line plots the mean of log Berk and Green model implied fund size for each portfolio. The blue line plots the mean of log fund size in the data for each portfolio. The purple dashed line plots the mean of log restricted model implied fund size for each portfolio. In top left figure, purple line plots fund size when there is no $\xi$. In top right figure, purple line plots fund size when there is no age. In bottom left figure, purple line plots fund size when there is no fund family size. In bottom right figure, purple line plots fund size when there is no marketing. The ten portfolios are formed on net skill. Portfolio 1 has the lowest net skill while portfolio 10 has the highest net skill.
Table A.18: Data Definition

<table>
<thead>
<tr>
<th>Variable</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fund AUM</td>
<td>Fund’s total net assets under management at the beginning of each year, in unit of millions of dollars</td>
</tr>
<tr>
<td>Fund expense ratio</td>
<td>The ratio between operating expenses that shareholders pay to the fund and the fund’s AUM</td>
</tr>
<tr>
<td>Actual 12b1</td>
<td>Reported as the ratio of the AUM attributed to marketing and distribution costs</td>
</tr>
<tr>
<td>Management fee</td>
<td>The ratio of the AUM attributed to fund management costs</td>
</tr>
<tr>
<td>Fund turnover</td>
<td>Minimum (of aggregated sales or aggregated purchases of securities), divided by the average 12-month AUM of the fund</td>
</tr>
<tr>
<td>Total market</td>
<td>Sum of all funds’ AUM including both active funds and index fund</td>
</tr>
<tr>
<td>Market share</td>
<td>Ratio between fund’s AUM and total market in the same year</td>
</tr>
<tr>
<td>Age</td>
<td>Number of years fund is in the sample prior to given year</td>
</tr>
<tr>
<td>Family size</td>
<td>Number of funds in the same fund family</td>
</tr>
<tr>
<td>CAPM $\alpha$</td>
<td>Outperformance estimated by CAPM</td>
</tr>
<tr>
<td>FF3 $\alpha$</td>
<td>Outperformance estimated by Fama French 3 factor model</td>
</tr>
<tr>
<td>FFC $\alpha$</td>
<td>Outperformance estimated by Fama French and Carhart model</td>
</tr>
<tr>
<td>FF5 $\alpha$</td>
<td>Outperformance estimated by Fama French 5 factor model</td>
</tr>
<tr>
<td>New</td>
<td>Dummy which equals 1 if fund is new in the current period</td>
</tr>
<tr>
<td>Index fund price</td>
<td>Fund expense ratio of the index fund</td>
</tr>
</tbody>
</table>

This table presents the data definition of all the variables used in the paper. For detailed data construction process, please check the data appendix.
Table A.19: Summary Statistics

<table>
<thead>
<tr>
<th>Percentiles</th>
<th>Num of Obs</th>
<th>Mean</th>
<th>Stdev</th>
<th>25%</th>
<th>50%</th>
<th>75%</th>
</tr>
</thead>
<tbody>
<tr>
<td>FF5 α (%)</td>
<td>27,621</td>
<td>0.54</td>
<td>7.98</td>
<td>-3.47</td>
<td>0.07</td>
<td>3.79</td>
</tr>
<tr>
<td>Fund AUM (million $)</td>
<td>27,621</td>
<td>1339</td>
<td>4791</td>
<td>82</td>
<td>254</td>
<td>886</td>
</tr>
<tr>
<td>Fund exp ratio (%)</td>
<td>27,621</td>
<td>1.66</td>
<td>0.53</td>
<td>1.23</td>
<td>1.75</td>
<td>2.05</td>
</tr>
<tr>
<td>Marketing expenses (%)</td>
<td>27,621</td>
<td>0.61</td>
<td>0.44</td>
<td>0.01</td>
<td>0.89</td>
<td>1.00</td>
</tr>
<tr>
<td>Market share</td>
<td>27,621</td>
<td>0.0018</td>
<td>0.0066</td>
<td>0.0001</td>
<td>0.0002</td>
<td>0.0009</td>
</tr>
<tr>
<td>Age</td>
<td>27,621</td>
<td>11.46</td>
<td>10.3</td>
<td>4</td>
<td>8</td>
<td>16</td>
</tr>
<tr>
<td>New dummy</td>
<td>27,621</td>
<td>0.0827</td>
<td>0.2755</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Family size</td>
<td>27,621</td>
<td>12.08</td>
<td>13.15</td>
<td>3</td>
<td>7</td>
<td>17</td>
</tr>
<tr>
<td>Index fund price (%)</td>
<td>27,621</td>
<td>0.17</td>
<td>0.09</td>
<td>0.13</td>
<td>0.17</td>
<td>0.19</td>
</tr>
<tr>
<td>Total market AUM (trillion $)</td>
<td>27,621</td>
<td>1.54</td>
<td>0.75</td>
<td>1.26</td>
<td>1.77</td>
<td>2.13</td>
</tr>
<tr>
<td>Family AUM (million $)</td>
<td>27,621</td>
<td>27,826</td>
<td>77,700</td>
<td>729</td>
<td>4,920</td>
<td>15,787</td>
</tr>
<tr>
<td>FFC α (%)</td>
<td>27,621</td>
<td>0.55</td>
<td>7.86</td>
<td>-3.41</td>
<td>0.25</td>
<td>3.97</td>
</tr>
<tr>
<td>FF3 α (%)</td>
<td>27,621</td>
<td>0.65</td>
<td>8.13</td>
<td>-3.39</td>
<td>0.25</td>
<td>4.09</td>
</tr>
<tr>
<td>CAPM α (%)</td>
<td>27,621</td>
<td>0.97</td>
<td>9.68</td>
<td>-3.76</td>
<td>0.45</td>
<td>4.92</td>
</tr>
</tbody>
</table>

This table presents summary statistics for our sample of U.S. equity mutual funds. For detailed variable definitions see table A.18. The sample period is from 1964 to 2015. Our unit of observation is fund/year.
Table A.20: Summary of Outcomes for Current Equilibrium and No-Marketing Equilibrium

|                                | Current | No-Marketing | No-Marketing  
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>(no foc errors)</td>
</tr>
<tr>
<td>Mean price (bp)</td>
<td>160.27</td>
<td>82.96</td>
<td>79.85</td>
</tr>
<tr>
<td>Mean marketing (bp)</td>
<td>61.29</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mean alpha (bp)</td>
<td>37.24</td>
<td>41.07</td>
<td>34.55</td>
</tr>
<tr>
<td>Total share of active funds</td>
<td>0.74</td>
<td>0.67</td>
<td>0.66</td>
</tr>
<tr>
<td>Mean sampling prob (%)</td>
<td>0.085</td>
<td>0.078</td>
<td>0.078</td>
</tr>
<tr>
<td>Sampling prob for low price funds (%)</td>
<td>0.042</td>
<td>0.14</td>
<td>0.20</td>
</tr>
<tr>
<td>Sampling prob for index fund (%)</td>
<td>5.91</td>
<td>13.66</td>
<td>13.66</td>
</tr>
<tr>
<td>Investor welfare (bp)</td>
<td>-140.72</td>
<td>-61.25</td>
<td>-66.26</td>
</tr>
<tr>
<td>Active funds average profits (bp)</td>
<td>57.51</td>
<td>42.19</td>
<td>48.45</td>
</tr>
<tr>
<td>Index fund profits (bp)</td>
<td>2.32</td>
<td>2.86</td>
<td>3.01</td>
</tr>
<tr>
<td>Total Welfare</td>
<td>-37.37</td>
<td>-16.20</td>
<td>-14.79</td>
</tr>
<tr>
<td>Investor’s Search Cost (bp)</td>
<td>29.09</td>
<td>12.15</td>
<td>10.48</td>
</tr>
</tbody>
</table>

This table provides various measures of the mutual fund industry under current and no marketing equilibrium. Additionally, we provide those measures for the no marketing equilibrium with no foc errors. Mean price, mean marketing and mean alpha are the arithmetic average of price, marketing expenses and alpha for all active funds, respectively. Total share of active funds is the market share of all active funds. The rest of the market share belongs to index fund. Sampling prob for low price funds is the mean sampling probability for the funds whose prices are below the mean price. Investor welfare is defined in equation 2.19. Active funds average profits is the mean of price minus marketing expenses for all active funds. Index fund profits is defined similarly. Total welfare is the sum of investor welfare, funds’ total profits and total marketing expenses. Investor’s search cost is the average total incurred search costs.
In this table, we provide the sensitivity of parameter estimates to two moment conditions: for fund pricing ($\sum_{t=1}^{T} \sum_{j \in t} \zeta_{j,t} = 0$) and marketing expenditures ($\sum_{t=1}^{T} \sum_{j \in t} \omega_{j,t} = 0$), for four subsamples that we use to estimate the model. “Interior” subsample includes only funds that are between the upper and the lower bounds on the marketing expenditures. “Lower” includes only funds whose marketing expenses are zero. “Upper” refers to the funds whose marketing expenses are at 100 bps. “All” refers to all of the funds in our sample. The magnitudes can be interpreted as the percentage bias of a parameter estimate if a moment is violated by 1% of the standard deviation of the corresponding moment condition error, $\omega$ or $\zeta$. For example, if the average $\zeta$ is not zero but instead equals to 1% of the estimated standard deviation of $\zeta$, then the estimate for $\lambda$ would be downward biased by 0.64% from its true value (if all funds are used).
BIBLIOGRAPHY


Bernstein, Shai, 2015, Does going public affect innovation?, *Journal of Finance* 70, 1365–1403.


Egan, Mark, 2017, Brokers vs. Retail Investors: Conflicting Interests and Dominated Products.,


Olson, Mancur, 1982, The rise and decline of nations: Economic growth, stagnation, and social rigidities, .


