Essays in Financial Economics

Zihan Huang
Pembroke College

A thesis submitted for the degree of
Doctor of Philosophy

Supervised by
Professor Martin Schmalz and Professor Mungo Wilson

Saïd Business School
University of Oxford

Trinity Term, 2026

Declaration

I declare that this thesis is entirely my own work and, except where stated, describes my own research.

Zihan Huang

Pembroke College

Acknowledgements

I am deeply grateful to my academic supervisors, Professor Martin Schmalz and Professor Mungo Wilson, for their unwavering support and guidance throughout my studies. I would also like to thank my family and my girlfriend for their constant encouragement and support over the years.

Abstract

This thesis consists of three essays in financial economics.

The first essay studies interlocking directorates among firms competing in the same product market. Using an exogenous shock that increased the litigation risk associated with interlocking directorates, I find that firms exposed to the shock experienced a significant decline in market value. I also find that the level of product market competition is negatively correlated with post-event stock returns. These findings support the theory that firms collaborate with competitors in product markets through shared board members.

The second essay investigates single-stock ETFs. I propose and provide evidence for the theory that these innovative products are primarily traded by attention-driven retail investors to circumvent leverage and short-selling constraints. I further show that single-stock ETFs lead to an increase in the idiosyncratic volatility of the underlying stocks.

The third essay studies the mutual fund industry. I propose a dynamic principal-agent model with private managerial skill and time-varying managerial effort. Consistent with the model's predictions, I provide empirical evidence that fund returns are not persistent and that fund size is an increasing but concave function of past returns.

Contents

Interlocking Directorate, Firm Value, and Product Market Competition

Abstract

I document a 500% increase in directorate interlocks with product market competitors among U.S. non-financial public firms over the past decade. Utilizing announcements from the U.S. Department of Justice regarding the enforcement of Section 8 of the Clayton Act as an exogenous shock to litigation risk associated with interlocks, I examine the net value of directorate interlocks with product market competitors to a firm. I validate the shock through both a difference-in-differences analysis and a regression discontinuity design using the threshold on equity value, demonstrating that Section 8 enforcement caused firms to lose directorate interlocks in the subsequent months. Then, I show that treated firms experienced a 7.1% decline in firm value within eight weeks following the announcement shock when compared to control firms while controlling for Fama-French five factors. Regression discontinuity analysis confirms the causal impact on stock prices. Moreover, the recent escalation in interlocking directorates was primarily attributed to the subset of firms exhibiting the highest product similarity with their competitors. Both product similarity with competitors and the number of competitors were negatively correlated with stock returns following the shock. These results support the theory that companies collaborate with competitors in the product market via shared board members. The quality of overlapping directors is not correlated with post-event stock returns, rejecting the notion that firms lose high-quality directors as a consequence of the shock.

1. Introduction

Corporate boards have been an active area of research in the field of corporate governance because of their significant influence on firm performance. Two major functions of a corporate board widely studied in the literature are the monitoring of firm managers and the advisory on firm strategy. Substantial evidence has been found that directors provide valuable service to the firm by providing their knowledge and expertise when managers set firm strategies. This unilateral information flow from the outside to the inside of the firm is one of the major motivations behind the setting of corporate boards in the first place. An implicit assumption behind this advisory function is that directors possess some expertise that they acquire outside of the firm, potentially from their professional service, including employment and directorship, at other firms. This assumption naturally gives rise to the possibility of bilateral information flow between the two firms, to which a director provides simultaneous service. Nonetheless, the role of bilateral information flow through shared directors has been underexplored in the literature, especially in the case of information exchange between competing firms, where the impact of such bilateral information flow on firm performance and market outcome may be the most influential.

An important question that remains unresolved in the literature is whether common directors between rival firms benefit shareholders by enabling two-way information exchange and encouraging anti-competitive behavior between the two firms. Despite its economic significance, this question has not been answered, potentially due to the scarcity of empirical data, as the United States has banned interlocking directorate, the appointment of the same director to two firms, between product market rivals since 1914, driven by anti-competitive concerns. In this paper, I document a sharp increase in the prevalence of interlocking directorate between competing firms in the past decade, suggesting a lack of enforcement of this prohibition in the recent decade. Such a rise in interlocking directorate at competing firms allows researchers to empirically examine the role of directors in the information flow between competing firms. To enable identification, I utilize a regulatory shock that leads to an increased probability of enforcement against interlocking directorate between competitors. Through an event study on this regulatory shock, I illustrate that firms gain benefits from directorate interlocks with their rivals. Subsequently, I demonstrate that these benefits are positively related to the intensity of competition the firm faces in the product market. The observed positive correlation between the intensity of competition and the benefits to firms substantiates the theory that companies coordinate in the product market via bilateral information flow through shared directors. This finding also refutes the alternative theory suggesting that interlocks increase firms' access to high-quality directors.

In 1914, the United States Congress enacted the Clayton Antitrust Act to enhance competition among American businesses. Section 8 of the Clayton Act prohibits competing companies in the product market from having directors in common. Despite being in force for over a hundred years, the enforcement of Section 8 of the Clayton Act has been largely neglected by regulators for a long time until 2022. During the week of April 4th, 2022, Jonathan Kanter, who leads the Antitrust Division at the US Department of Justice, made two public announcements regarding the DOJ's plan to begin enforcing Section 8 of the Clayton Act. This week signified the commencement of Section 8 enforcement by several US regulators. Before April 2022, interlocking directorates among US non-financial public companies had been rapidly increasing since the early 2010s. In figure 1, I document both the decade-long rise in interlocking directorates and the sharp decline following the DOJ announcements. Both panels in figure 1 illustrate this upward trend in directorate interlocks among US firms over the past decade. In the first panel, I display the average number of interlocked firms, defined as those that share at least one common director with the focal firm, across all

US non-financial public firms. This number rose by approximately 20% from the early 2010s low point to the peak in 2022. The second panel shows an enormous 500% rise in the average number of interlocked product market competitors across all firms from the early 2010s to 2022. The annual jumps in the average number of interlocked competitors in the second panel are mechanical due to the use of the annual competitor metric from Hoberg and Phillips (2016). For each month, I use the Hoberg and Phillips (2016) measure from the previous year. This practice is consistent with the analysis in the following sections. If I use the competitor metric in the concurrent year, there is a discrete drop in the average number of interlocked competitors at the beginning of each year. The firms analyzed in this paper are all US non-financial public firms, as the Clayton Act does not apply to the finance industry. While the increase in the number of US non-financial public firms may partially explain the patterns in the first panel, the consistent and substantial rise shown in the second panel cannot be merely attributed to the growth in the number of public firms. Crucially, since the Clayton Act Section 8 aims to prevent interlocks among product market competitors, the significant increase in the second panel highlights the lack of enforcement of Clayton Act Section 8 before 2022. Another key observation from figure 1 is the sharp decline in the number of interlocks starting in April 2022, marked by the vertical dashed lines. This serves as the first piece of evidence for the effectiveness of regulators' efforts to prohibit interlocking directorates following the DOJ announcements.

To address the research questions concerning whether and how firms benefit from interlocking directorates with competitors, I perform empirical analysis using the previously mentioned DOJ announcements as an exogenous shock on the likelihood of enforcing interlocking directorate prohibitions. My analyses are guided by 2 hypotheses. The first hypothesis is that:

1. The enforcement of prohibition against interlocking directorate will induce firms to sever their interlocking directorate links with other firms, especially their product market competitors.

If the interlocking directorate somehow benefits the interlocked firms, hypothesis 1 will naturally lead to the following second hypothesis:

2. Following the news of Section 8 enforcement, we should see a drop in stock price of firms interlocked with competitors, relative to the price of firms without competitor interlocks.

I design empirical tests for each of the hypotheses mentioned above. To more precisely determine which companies are significantly impacted by Section 8 enforcement, I identify the treated companies based on three criteria: board overlap, product market competition, and book value of equity. These criteria are explicitly stated by the Clayton Act Section 8 to determine if a firm is subject to the prohibition outlined in the Clayton Act Section 8.

After identifying treated firms, I perform an event study examining the number of interlocks around DOJ announcements to test hypothesis 1. At the firm-month level, I determine the number of firms that shared at least one director with the focal firm. By conducting a difference-in-differences analysis, I find that on average, a treated firm lost 0.203 more interlocked firms compared to a control firm over six months after the announcements. After restricting the dependent variable to include only interlocked firms with a sufficient book value of equity to be targeted by Section 8, the estimated treatment effect expands to -0.377. After narrowing the variable to encompass solely the interlocked competitors in the product market that had substantial equity values, the estimated treatment effect stays at -0.352. The treatment effects estimated for all three variables are not only statistically significant but also economically significant. Prior to the treatment, treated firms had an average of 4.09 interlocked above-threshold competitors. The treatment effect represents nearly 10% of these interlocks. Furthermore, the speed at which the treatment proved effective was noteworthy. Event study graphs show that the majority of the treatment effect took place within two months after the treatment.

To enhance causal interpretation, I employ a sharp regression discontinuity design by leveraging the sharp threshold on firm book value of equity. Section 8 does not apply to a pair of firms if either firm had a book value of equity below the threshold. Regression discontinuity analysis demonstrates that, following DOJ announcements, the firms just above the equity threshold experienced more interlock losses compared to those just below it. The local treatment effect identified in this analysis also grows over time, aligning with the pattern shown in the event study. The findings from the event study and regression discontinuity both validate the policy shock discussed in this paper, and these findings endorse hypothesis 1.

After confirming hypothesis 1, I proceed to test the second hypothesis by performing an event study on stock returns around the announcement events. I observe that in the week after the event, firms that were treated experienced lower stock returns than the control firms. On average, the stock prices of the treated firms underperformed control firms by 7.1%, after accounting for the Fama-French 5 factors, during the 8 weeks following the announcement events. This underperformance is both statistically and economically significant. These findings indicate that the net value of a directorate interlock with a competitor is significantly positive for the firms in my sample. I demonstrate that the poor performance of the treated firms was not driven by variations in stock performance across different industries or differential exposure to inflation risks.

I perform a regression discontinuity analysis on stock returns to demonstrate the causal relation between announcements on Section 8 enforcement and the ensuing negative reaction in the stock market. Regression discontinuity analysis shows that firms just above the equity threshold underperformed the firms just below the threshold, after accounting for Fama-French 5-factors, on the trading day following the announcement. By leveraging the efficiency of the stock market, I provide evidence that firms enjoy net benefits from having interlocks with competitors after taking all costs and benefits into account. This finding provides insight into the debate on whether interlocking directorate is good or bad for the firms (Geng, Hau, Michaely, and Nguyen, 2026; Gopalan, Li, and Zaldokas, 2026; Cabezon and Hoberg, 2026; Begley, Haslag, and Weagley, 2026; Herrera-Caicedo, Jeffers, and Prager, 2024; Barone, Schivardi, and Sette, 2025; Geng, Hau, Michaely, and Nguyen, 2025).

Lastly, I perform heterogeneity analyses to distinguish between different channels through which competing firms might benefit from directorate interlocks. First, I demonstrate that the recent rise in interlocking directorates among U.S. public companies is largely due to the subset of firms that have the greatest product similarity with their competitors. Secondly, I conduct a difference-in-differences analysis with triple interaction terms that include the TNIC-4 total similarity score and the number of TNIC-4 level competitors. The results indicate that treated firms with higher total similarity to competitors and a higher number of competitors performed worse than other treated firms. This indicates that companies with closer competitors benefit more from interlocks with those competitors. Lastly, I directly test the alternative hypothesis of lower director quality following the regulation enforcement. Using firm-level measures that are correlated with the quality of the overlapping directors, I find that the post-event decrease in the stock prices of treated firms is not correlated with the ex-ante overlapping director quality.

The findings in the heterogeneity analysis support the theory of collusion in the product market via shared directors. The product market collusion channel can simultaneously explain these findings that firms with more analogous competitors are more likely to have directorate interlocks and that these firms derive greater value from such interlocks. Canonical economic models indicate that when rival companies offer more similar products, they tend to earn lower profits as a result of price competition. Consequently, it would be mutually beneficial for these companies to agree on collusion. In practice, two main factors restrict firms from engaging in collusion, yet these can be alleviated through interlocking directorates. First, if collusive coordination is uncovered, the government will impose penalties. Nonetheless, the private setting of board meetings permits shared directors to secretly coordinate among rival firms without drawing the attention of regulators. Second, mutual trust is essential between colluding firms; without it, each party is tempted to deviate. This trust can be established if a shared director consistently oversees the product market strategies of both firms and acts as a guarantor of their adherence. Thus, the empirical results observed in the heterogeneity analysis are natural consequences of profit-maximizing firms appointing shared directors with competitors to facilitate collusion and achieve greater profits.

The findings from the heterogeneity analysis also reject the alternative theory of lower board quality, which suggests that the decrease in firm value following the DOJ announcements is due to the possible exit of high-quality directors from affected companies. The logic behind this theory is that after announcements by the DOJ, a director can now serve simultaneously at fewer competing firms, reducing the overall pool of qualified directors available. In equilibrium, firms hire the most competent directors. Therefore, following the enforcement of Section 8, many affected companies will have to replace high-quality directors with lower-quality ones, resulting in diminished monitoring and advisory benefits for the firm and ultimately lower firm values. The results from the heterogeneity analysis reject this board quality theory because, if the board quality channel is true, the expected reduction in monitoring and advisory values following Section 8 enforcement should be correlated with the ex-ante quality of overlapping directors, which is empirically rejected.

The heterogeneity analysis reveals that a key role of the corporate board is to serve as a discreet means of communication and coordination among rival companies. The significance of this function is on the rise, given that the practice of having directors shared among competing companies has become more prevalent over the past decade. Unlike the typical advisory role, characterized by a one-way flow of information from external sources into the firm, this communication function involves a bilateral exchange of information between rival firms and might have more substantial implications for stakeholder welfare.

The remainder of the paper proceeds as follows. Section 2 discusses relevant literature and my contributions. Section 3 provides background on the Clayton Act Section 8 and details relevant to my analyses. Section 4 introduces data. Section 5 presents empirical analysis. Section 6 concludes the paper.

2. Related Literature and Contributions

This paper contributes to two strands of literature. Firstly, it contributes to the corporate governance literature concerning the role of boards. The literature on corporate governance has identified two primary roles of corporate boards (Adams and Ferreira, 2007; Adams, Hermalin, and Weisbach, 2010). The first function is monitoring the top management team (Hermalin and Weisbach, 1998; Adams and Ferreira, 2009), and the second function is advising on the firm strategy (Coles, Daniel, and Naveen, 2008; Dass, Kini, Nanda, Onal, and Wang, 2014). Research on board advisory has shown that the industry and professional expertise of outside directors contributes to increasing firm value (Dass et al., 2014; Drobetz, von Meyerinck, Oesch, and Schmid, 2018). However, this body of literature inherently concentrates on the one-way flow of information from outside the firm to inside. I contribute to this body of literature by demonstrating that a board can enhance a firm's value by serving as a communication channel and enabling bilateral information flow between companies. My focus is on the bilateral information exchanges between competing firms, where such exchanges can be particularly advantageous for the firms involved. I first demonstrate the net positive value generated by competitor interlocks, and then I present evidence supporting the theory that such value enhancement is achieved through bilateral information flow between firms.

I also provide insights to the most relevant contemporaneous literature on direct board interlocks between product market competitors. This strand of literature has provided evidence that direct interlocks between competitors could be both beneficial and costly to the firm. Such interlocks could benefit the firm through higher product prices (Barone et al., 2025; Gopalan et al., 2026), lower expenses (Geng et al., 2026), and more segmented markets (Poberejsky, 2024). At the same time, competitor interlocks could lead to innovation herding, which harms competitive edge and long-term profits (Cabezon and Hoberg, 2026). In this paper, I estimate the aggregate firm value, after taking all potential channels into account, created by competitor interlocks by examining stock prices. I find that directorate interlocks positively affect the value of a firm. This finding informs the current debate on the impact of competitor interlocks on firm value.

3. Institutional Background

Among the US antitrust statutes, Clayton Antitrust Act has been one of the primary building blocks since it was first enacted in 1914[1]. In 1914, the United States Congress passed the Clayton Antitrust Act that aimed at modifying and complementing the then-existing federal antitrust laws to further facilitate competition among US businesses. The Clayton Act targets many business practices that may harm competition. Of particular interest for this paper, Section 8 of the Clayton Act prohibited "interlocking directorates" among competing non-financial firms, except for firms that have less than 1 million dollars in "capital, surplus, and undivided profits," which are legal terms that map to components of firm equity[2]. Importantly, the Clayton Section 8 prohibition is *per se*, meaning that the act of having an interlocking directorate, by itself, is a violation of antitrust law, no matter whether it has an ultimate impact on the competition between the interlocked firms[3].

The interlocking directorate ban by the Clayton Act was not uncontroversial, especially around the end of the 20th century. The advocates for a reform of the Clayton Act claimed that Section 8 prevented firms from hiring qualified directors, even when such hiring would not lessen competition. As a result, the US Congress amended the Clayton Act in 1990. The major modifications to the Clayton Act include extending the scope of Section 8 from directors to both directors and officers, adding several safe harbor exceptions to Section 8, modifying the equity threshold, and charging the Federal Trade Commission to adjust this threshold annually based on Gross National Product[4]. Among these modifications, the most relevant to this paper is that the FTC

[1] Refer to "https://www.govinfo.gov/content/pkg/COMPS-3049/pdf/COMPS-3049.pdf" for the latest version of the Clayton Act

[2] https://www.law.cornell.edu/supremecourt/text/275/215/USSC_PRO_275_215_156

[3] https://www.justice.gov/opa/pr/justice-department-s-ongoing-section-8-enforcement-prevents-more-potentially-illegal

[4] https://www.congress.gov/bill/101st-congress/house-bill/29

is demanded by Congress to publish an updated equity threshold each year. As of 2022, the focal year of this paper, the latest equity threshold was $41,034,000[5]. This means that in 2022, all firms with a book value of equity greater than this amount would be subject to the interlocking directorate ban of the Clayton Act Section 8, while all firms below this threshold were exempted. The amendment also added safe harbor exemptions for firm pairs that are not significant "competitors"[6]. The amendment stipulates that a firm pair is exempted from Section 8, as long as any one of the three conditions listed below is satisfied. The first condition is that the competitive sales of either firm in the pair are below a sales threshold. The second condition is that the competitive sales of either firm in the pair are below 2% of its total sales. The third condition is that the competitive sales of each firm in the pair are below 4% of its total sales, respectively.

Since the 1990 reform of the Clayton Act, the enforcement of Section 8 had been slack, justified by the rare violations of Section 8 before 2014[7]. However, the interlocking directorate between product market competitors started to rise since 2014, as shown in figure 1. Despite such a sharp increase in board interlocks between competitors, the enforcement of Section 8 did not keep up with the rise in interlock. Then, this situation of under-enforcement changed in the week of April 4th, 2022, when two significant regulatory events took place. At the 2022 Spring Enforcers Summit on April 4th, 2022, the assistant attorney general Jonathan Kanter of the US Department of Justice Antitrust Division delivered a speech in which he announced a step-up in the Clayton Act Section 8 enforcement[8]. Then on April 8th, 2022, Friday of the same week, Kanter attended the roundtable event at the American Bar Association Antitrust Law

[5]https://www.ftc.gov/news-events/news/press-releases/2022/01/ftc-announces-annual-update-size-transaction-thresholds-premerger-notification-filings-interlocking

[6]https://www.congress.gov/bill/101st-congress/house-bill/29

[7]https://www.justice.gov/archives/opa/speech/remarks-assistant-attorney-general-bill-baer-american-bar-association-clayton-act-100th

[8]https://www.justice.gov/opa/speech/assistant-attorney-general-jonathan-kanter-delivers-opening-remarks-2022-spring-enforcers

Section spring meeting and reiterated DOJ's intent to step up Section 8 enforcement[9]. These announcements were surprises to the public, and they were the first high-profile indications of DOJ's intent to enforce Section 8. While the Spring Enforcers Summit was mainly attended by government officials around the world and may not attract much attention from the industry and the market, the ABA meeting was targeting all practicing lawyers in the United States. Therefore, it's reasonable that these two announcements together would affect both firm management teams and the stock market. Since this week of announcements, DOJ has taken intense regulatory and legal actions against interlocking directorate. In the following months, numerous firms severed their director links with each other[10]. DOJ was later joined by other regulators, such as the Federal Trade Commission, in the effort to enforce Section 8[11]. It's obvious that the current wave of Section 8 enforcement has made huge impacts on US firms, and all of these started with the two announcements by Kanter in the week of April 4th, 2022. Figure 1 validates this claim by showing that interlocking directorate among US public firms started to decrease since April 2022, reversing the decade-long trends. Given that the Clayton Act was passed but rarely enforced for decades, these announcements were largely unanticipated by the market. Because these two significant events took place in the same week, I use this week as the event week in my following analysis.

The Department of Justice's announcements in April 2022 were a component of the Biden administration's broader initiative to enhance competition in the United States. This competition-promoting initiative commenced with an executive order in July 2021[12]. In the following months, multiple government agencies intensified antitrust

[9] https://www.justice.gov/archives/opa/video/assistant-attorney-general-jonathan-kanter-virtually-participated-enforcers-roundtable-aba

[10] https://www.justice.gov/opa/pr/directors-resign-boards-five-companies-response-justice-department-concerns-about-potentially

[11] https://www.ft.com/content/33f0d653-c53b-4844-a044-3f72507a518b

[12] https://bidenwhitehouse.archives.gov/briefing-room/presidential-actions/2021/07/09/executive-order-on-promoting-competition-in-the-american-economy/

enforcement in traditional areas, such as merger and acquisition and labor market competition. In April 2022, antitrust enforcement was broadened to include interlocking directorates, an area that had previously seen limited enforcement. This antitrust initiative stemmed from Biden's narrow victory in the 2020 election, suggesting that the government's antitrust efforts were likely exogenous to the interlocking directorates among U.S. firms. Additionally, due to the fact that the enforcement announcements for Clayton Act Section 8 followed years of inadequate enforcement of this particular section, it is probable that neither companies nor investors anticipated this enforcement against interlocking directorates, even in light of Biden's antitrust efforts.

There are two nuances in the Clayton Act that are worth mentioning. First, the 1990 amendment of the Clayton Act extended the scope of interlocks from only directors to both directors and senior officers. Therefore, in this paper, my definition of interlocking directorate includes all directors on the board and C-suite executives. However, the inclusion of senior executives makes a tiny difference, as detailed in the next section, in the main variables and the treatment classification, and all results would remain qualitatively similar if I excluded senior executives from consideration. This shows that interlocked non-director executives are very rare among US non-financial public firms. For the sake of brevity, I refer to directors and senior executives simply as "directors" in the remainder of this paper. Second, simultaneous board positions at two firms by different natural persons representing the same investment institutions are also prohibited[13]. I cannot empirically account for this phenomenon because I do not have access to affiliations of directors to investment institutions. However, this should not be concerning, because institutional investors rarely hold simultaneous boards in competing public firms (Geng et al., 2025).

[13]https://www.ftc.gov/enforcement/competition-matters/2017/01/have-plan-comply-bar-horizontal-interlocks

4. Data

The data used in this paper come from 4 sources. First, I obtain data on the board of directors and C-suite executives (including CEO, CFO, and COO) from BoardEx for all US public firms. Only the CEO, CFO, and COO are considered because the Clayton Act Section 8 defines an "officer" as any executive who is appointed by the board of directors. BoardEx specifies the basic information for each board member and executive, and it assigns a director ID to each individual who has ever appeared in the dataset. With this information, I could identify the directors and C-suite executives who have served simultaneously at different companies. I include C-suite executives in the data because the 1990 Amendment of the Clayton Act Section 8 extends the scope from directors to both directors and officers, which are typically defined as board-selected senior executives in legal interpretations. Nonetheless, the incidence of non-director senior executives simultaneously sitting at another public firm is extremely rare. Including senior executives would result in a 3% increase in the average number of both interlocked firms and interlocked competitors and a 2% increase in the number of treated firms. These differences are small, and repeating all analyses using director-only interlocks makes no significant difference in the results. For this reason, I refer to the interlocks of both directors and senior executives as director interlocks in this paper. Second, I obtain the stock prices and returns for all US-listed firms from CRSP. All stock returns reported in this paper are in decimal. Third, I obtain quarterly financial data for the firms from Compustat. Then, I link the firms in the BoardEx dataset with those in the CRSP and Compustat datasets using the linking tool provided by Wharton Research Data Services. I obtain SIC codes from CRSP and then augment them with Compustat data. Since the Clayton Act Section 8 only targets non-financial firms, I exclude financial firms from all following analyses. Following Cabezon and Hoberg (2026), I identify financial firms as those with SIC codes from 6000 to 6999. I compute equity as $cstkq + capsq + req$, closely corresponding to the terms "capital, surplus, and undivided profits". Lastly, I use Hoberg and Phillips (2016, 2010) text-based product competition data from the Hoberg and Phillips Data Library[14]. This data provides an annual metric that measures the pairwise product market competition between all US public firms, based on a textual analysis of their annual financial reports. For all firm pairs, I use the TNIC scores from the fiscal year 2021, which are computed from firms' annual reports for the fiscal year 2021. Because the fiscal years of the majority of US public firms conclude around the end of the calendar year, the fiscal year 2021 annual reports for most firms are released in the early months of 2022. Given that the DOJ announcement events took place in April, the most up-to-date similarity metrics the market could observe in the event week are for the fiscal year 2021.

5. Empirical Analysis

The empirical strategy of this paper consists of three parts. First, I establish a relation between the enforcement of the Clayton Act Section 8 and the reduction in firms' interlocking directorate connections. To do so, I deploy a difference-in-differences analysis to investigate whether firms reduced their directorate interlocks following the DOJ announcements and whether there was a significant trend before the announcement events. This exercise tests the first hypothesis discussed earlier. Second, I investigate whether firms affected by the Clayton Act Section 8 suffered from lower stock returns than other firms following the DOJ announcements by conducting an event study around the week of the announcements. This provides an estimate of the overall firm value attributed to directorate interlocks with product market competitors. Third, I conduct a heterogeneity analysis on the interlocking directorates and stock returns to

[14] http://hobergphillips.tuck.dartmouth.edu

shed light on the channel through which directorate interlocks could contribute to the firm value.

5.1. Definition of Firm-Level Treatment

In this paper, I conduct two event studies to investigate the differences in both interlocking directorate and stock returns between firms affected by the Clayton Act Section 8 and firms that were not affected. Before running difference-in-differences analysis, I first need to define the treated and control firms. To define the treatment group, I rely on three criteria adapted from the Clayton Act Section 8. The Clayton Act Section 8 specifies that interlocking directorate prohibition is applicable to firm pairs that are both above the equity threshold, share a common director, and compete in the product market simultaneously. Therefore, I begin by identifying all firm pairs that satisfy all three criteria. First, I identify firm pairs as product market competitors if their annual TNIC score (Hoberg and Phillips, 2016, 2010) was in the top 1% among all firm pairs in 2021. This definition of product market competitor is identical to that in Cabezon and Hoberg (2026), who find that firm pairs within the top 1% of TNIC score are direct competitors in the product market and that interlocking directorate between these firm pairs results in innovation herding. According to these authors, the 1% threshold yields a competitor classification that is as coarse as the 4-digit standard industrial classification code. They show that this granularity better captures competition than does TNIC-3 level classification, which is likely to include complementary firms in the product market.

While the 1990 amendment of the Clayton Act Section 8 exempts non-competing firm pairs based on competitive sales revenues, these numbers are not easily accessible, so I follow recent literature (Cabezon and Hoberg, 2026; Hoberg and Phillips, 2016; Gopalan et al., 2026; Poberejsky, 2024; Begley et al., 2026) to use TNIC scores to identify product market competitors. It's reasonable to use this annually-updated text-based industry classification to proximate the threshold used by regulators because both competitive revenue and TNIC scores are closely related to the actual sales of products each firm had in the relevant year. By legal requirements, the products that constitute larger shares of revenue for a firm are more likely to be mentioned in the product description section in the annual report, from which the TNIC score is constructed each year. Nonetheless, it's important to keep in mind that the TNIC measure for competition is not likely to coincide with the measure used by enforcement agencies and that SIC 4-digit granularity is also likely to deviate from that adopted by the government.

Then, I further identify firm pairs with at least 1 overlapping director on March 31st, 2022, using BoardEx data. Lastly, I examine whether the latest book value of equity of both firms in the firm pair was greater than the equity threshold as of event week. If a firm pair satisfies all three criteria simultaneously, I identify this pair as a firm pair treated by the enforcement of the Clayton Act Section 8. To convert the firm pair-level treatment to firm-level treatment, I require a firm to be in at least one treated firm pair for the firm to be classified as a treated firm. This treatment classification is used in both event studies on interlocking directorate and stock returns. The summary statistics on main variables for both treated and control firms are presented in table 1.

5.2. Event Study on Number of Interlocks

After defining the firm-level treatment, I conduct the monthly level event study and difference-in-differences analysis on the number of directorate interlocks. I include all non-financial firms with non-missing book value of equity observations as of the end of March 2022 from Compustat. I require non-missing book value of equity because this value is necessary to conduct treatment classification.

5.2.1. *Event Study Plots on Number of Interlocks*

I first visualize the evolution of the differences in the number of interlocks between treated and control firms. I look at 3 dependent variables on the number of interlocks in this event study exercise. The first dependent variable is the number of non-financial firms that a firm shared any common director with. Note that not all common directors are prohibited by Section 8, as discussed earlier. Therefore, to further refine the measure, I also compute the number of interlocked firms that had the latest book value of equity greater than the 2022 threshold. This variable takes into account the equity value of the interlocked firm, and thus the interlocks included in this variable are more affected by Section 8 enforcement than the first variable. The third dependent variable, which is the most related to 3 criteria of Section 8, is the number of above-threshold non-financial competitors that a firm shared any common director with. This variable takes all three criteria of Section 8 into account. I plot the coefficients from the following regression equation:

$$Interlocks_{it} = \sum_{r=-6}^{6} \beta_r \times \mathbf{1}[Treated_i = 1] \times \mathbf{1}[t = r] \times \mathbf{1}[r \neq -1] + \delta_t + \lambda_i + \epsilon_{it} \quad (1)$$

where $Interlocks_{it}$ is the number of interlocks, as defined above, for firm i at the end of month t, which is relative to the event month of April 2022. $Treated_i$ is the treated dummy for firm i as defined earlier. The point estimates and standard errors for β_r are plotted in figure 2. We can see that following April 2022, treated firms experienced larger drops in all three measures of directorate interlocks. These results are consistent with hypothesis 1 that treated firms reduced directorate interlocks to lower the Section 8 enforcement risk. Furthermore, there were limited pre-trends for all three variables.

5.2.2. Difference-in-Differences on Number of Interlocks

After visualizing the change in the number of interlocks, I further conduct a difference-in-differences analysis to quantify the relation. The regression equation is as follows:

$$Interlocks_{it} = \alpha + \beta_1 \times Post_t \times Treated_i + \delta_t + \lambda_i + \epsilon_{it} \tag{2}$$

where $Interlocks_{it}$ is one of the three interlock variables defined above for firm i in week t. Treated dummy for firm i is the same as defined in section 5.1. I define post dummy as equal to 1 if the observation month is equal to or later than April 2022. I include firm and month fixed effects and double-cluster standard errors at firm and month level.

The results from this regression are presented in table 2. One can see that the coefficients of interaction terms are significantly negative across all three dependent variables. This suggests that treated firms experienced a larger reduction in interlocking directorates than control firms. The treated firms lost 0.203 more interlocked firms than control firms. This number includes interlocks with both firms that were targeted by Section 8 and firms that were too small to be targeted. When we restrict our interlock to those with firms that were large enough to be targeted by Section 8, the treatment effect magnifies to -0.377, which is also more statistically significant. The third column shows that treated firms reduce the number of interlocked above-threshold competitors by 0.352 more than control firms. The statistical significance of this estimate is also at 1%. Although the estimated value is slightly smaller for interlocked above-threshold competitors than for interlocked above-threshold firms, it represents a much larger percentage decrease in the number of interlocked above-threshold competitors, as the pre-event number of interlocked above-threshold competitors is more than 30% smaller than the pre-event number of interlocked above-threshold firms, as shown in table 1. These patterns suggest that firms severed Section 8-eligible interlocks more than Section 8-ineligible interlocks, supporting the hypothesis that firms severed those interlocks to lower the enforcement risk related to the Clayton Act Section 8. This event study on the number of interlocks suggests that regulators' efforts since the DOJ announcements in April 2022 are effective in reducing interlocking directorate among targeted firms, further validating the policy shock.

5.2.3. Sharp Regression Discontinuity on Number of Interlocks

To enhance the causal interpretation, I use a sharp regression discontinuity design to answer the question of whether Section 8 enforcement caused firms to reduce their directorate interlocks. To achieve this, I leverage the sharp equity threshold in Section 8. The Clayton Act Section 8 exempts all firm pairs from interlocking directorate prohibition if either firm has "capital, surplus, and undivided profits" smaller than an annually updated threshold. I obtain the latest quarterly book value of equity ($cstkq + capsq + req$) as of the event week from Compustat. The sharp equity threshold was $41,034,000 for the year 2022, as published by the FTC in January 2022.

A complication in the identification is that not every firm exceeding the threshold had an interlocking directorate with other companies. Consequently, some firms above the threshold were not impacted by Section 8 because they didn't have a qualified interlock. This suggests that the treatment probability does not jump from 0 to 100% at the equity threshold. In order to apply sharp regression discontinuity design, which requires a discrete jump in treatment probability from 0 to 100%, I restrict the sample of the firms in the regression discontinuity analysis such that all sample firms above the equity threshold were involved in at least one firm pair that was targeted by Section 8. In this section, I adopt two different subsamples because the exact competition measure and granularity adopted by enforcement agencies are not observable. In the first specification, I restrict my sample to the firms that had at least one interlocked firm that had an equity value above the threshold as of the last day in March 2022. In this specification, I do not consider the competition metric. In the second specification, I restrict my sample to the firms that had at least one interlocked competitor that had an equity value above the threshold. In this specification, I am conservative by incorporating the competition metric at TNIC-4 level. For each subsample, I estimate the following regression:

$$Y_i = \alpha + \beta_1 \times D_i + \beta_2 \times Equity_i + \beta_3 \times D_i \times Equity_i + \epsilon_i \tag{3}$$

where D_i is a dummy equal to 1 for firm i if its latest book value of equity was greater than \$41,034,000 as of the event week of April 4th, 2022. To facilitate interpretation, $Equity_i$ here is the latest book value of equity of firm i minus the equity threshold, in millions of dollars. The local treatment effect of interest will be captured by the estimate for β_1. Two dependent variables corresponding to two different subsamples are used. The first dependent variable is the change in the number of interlocked above-threshold firms. The second dependent variable is the change in the number of interlocked above-threshold competitors at TNIC-4 level. I compute the dependent variables at two horizons to capture the change over time. First, I compute the change from March 31st, 2022 to April 30th, 2022. This measures the change that happened within one month after the announcements. Second, I also compute the change from March 31st, 2022 to May 31st, 2022. This measures the change over the two months following the shock. Since appointing and removing directors requires time-consuming administrative processes, one may expect the change in interlocking directorate to gradually happen in a few months, as demonstrated in figure 2.

Note that the construction of this regression discontinuity analysis is not cross-sectional, as typical in regression discontinuity designs. It's similar to the "first-difference estimator" (Lemieux and Milligan, 2008), which is derived from taking the difference between two cross-sectional regressions. The intuition behind this construction is that the correct measure for the treatment effect is the difference in the outcome variable before and after the treatment. Lemieux and Milligan (2008) argue that this "first-difference estimator" is more stringent than cross-sectional regression discontinuity because it further controls for individual-specific fixed effects. This relaxes the assumption of cross-sectional regression discontinuity that individuals around the threshold are similar by comparing the changes within each individual at different times. These authors also address the potential selection bias concern in subsample regression discontinuity by proving that such selection bias will not affect the validity of this test as long as the selection bias is a smooth function of the running variable. There is no evidence for the violation of this condition in my sample.

I estimate the above regression with a bias-corrected estimator and the corresponding robust standard error (Calonico, Cattaneo, and Titiunik, 2014). I adopt a data-driven bandwidth selector to optimize the mean squared error (Calonico et al., 2014). Before applying the bandwidth selector, I manually drop firms with book values of equity that were either greater than 101 times the equity threshold or smaller than negative 99 times the equity threshold. This manual exclusion is very generous and unlikely to affect the estimation of treatment effects. Such manual exclusion on both sides is necessary for the above estimator to work. I report the bias-corrected point estimates for the treated effect and the corresponding robust standard errors in table 3.

From table 3, the first observation we can make is that the estimated local treatment effect is negative across all four specifications. By comparing results at different horizons in table 3, we can see that the treatment effect is increasing over time. This finding is consistent with the expectation that the change in directors should happen over a few months. Out of four specifications, the estimated local treatment effect is negative and significant in three specifications. The results in table 3 indicate that the DOJ announcements on Section 8 enforcement caused firms to sever their directorate interlocks. Overall, the results from difference-in-differences and regression discontinuity exercises in this section validate the DOJ announcements in April 2022 as a policy shock to the enforcement probability on interlocking directorate among US firms by showing that firms lost directorate interlocks following the announcements.

5.3. Event Study on Stock Returns

Having confirmed hypothesis 1 by showing the negative impact of the DOJ announcements on the number of director interlocks, I continue to test hypothesis 2 by investigating whether shareholders experienced a loss in stock values following the enforcement announcements. Academic literature (Barone et al., 2025; Geng et al., 2026; Gopalan et al., 2026; Herrera-Caicedo et al., 2024; Begley et al., 2026) has shown that reducing interlocking directorate benefits consumers and employees. However, it's crucial to understand the net effect of interlocking directorate prohibition on firm value to comprehensively understand the value of directorate interlocks with competitors. In the remainder of this section, I examine how stock prices evolved around the announcement events.

In this section, I first conduct an event study to investigate the stock market reaction around the events of DOJ announcements in the week of April 4th, 2022, and then I perform a heterogeneity analysis to investigate what firm characteristics are related to such reactions. As described earlier, the assistant attorney general Jonathan Kanter of the US Department of Justice Antitrust Division made two public announcements, indicating DOJ's interest in enforcing the Clayton Act Section 8, at both the 2022 Spring Enforcers Summit on April 4th, 2022, and the American Bar Association spring meeting on April 8th, 2022. While the Spring Enforcers Summit was mainly attended by regulators around the globe, the ABA spring meeting was targeting all practicing lawyers in the US. These pieces of news were largely unanticipated by the market because they were the first, in the recent decades, high-profile indication of DOJ's intent to enforce Section 8 and to crack down on interlocking directorate. It signaled a shift in the emphasis of the DOJ Antitrust Division, and then more US regulators followed suit. Since both announcements were made in the same week, I use this week as the event week and conduct the event study at the weekly level.

5.3.1. Weekly Event Study Plot on Stock Returns

To capture the difference in the stock returns between treatment and control groups, I run the following two-way fixed effect regression:

$$R_{it} = \sum_{r=-8}^{8} \beta_r \times \mathbf{1}[Treated_i = 1] \times \mathbf{1}[t = r] \times \mathbf{1}[r \neq -1] + \delta_t + \lambda_i + \epsilon_{it} \tag{4}$$

where t is the week relative to the event week, which is the week of April 4th, 2022 in this case. $\mathbf{1}[Treated_i = 1]$ is an indicator variable that is equal to 1 if the firm i is treated by Section 8 enforcement, as defined in previous sections. I set both the pre-event window and post-event window to 8 weeks. Thus, r will range from -8 to 8. The dependent variable, R_{it}, is the week-end Fama-French 5-factor and 3-factor cumulative abnormal return, starting from the beginning of my event window. To compute cumulative abnormal return, I first estimate the Fama-French 5-factor model on daily stock returns of each stock using all daily observations in 2021, the year before DOJ announcements. Then I compute the daily Fama-French 5-factor and 3-factor abnormal returns for each stock on each day in my event window, and then aggregate daily abnormal returns to weekly abnormal returns. Finally, I use these weekly abnormal returns to compute weekly cumulative abnormal returns from the beginning of the event window. Fama-French abnormal returns are used to control for the potential confounding factors of different market exposure, size, growth perspective, profitability, and investment aggressiveness between treated and control firms. Coefficients of interest are β_r, which denotes the difference in return between treated firms and control firms in week r. β_{-1} is not estimated to avoid the multicollinearity problem. Standard errors are double clustered at the firm and week level, and fixed effects on firm and week are included. The point estimates and 95% confidence intervals for β_r in the above regression are plotted in figure 3.

In figure 3, differences in weekly Fama-French 5-factor and 3-factor cumulative abnormal returns between treated and control groups are plotted. Figure 3 shows that differences in both Fama-French 5-factor and 3-factor cumulative abnormal returns show very similar patterns around the announcement week. There are roughly parallel trends in the returns of treated and control firms in the pre-event window. Immediately after the event week, treated firms started to underperform control firms and continued to underperform in the following weeks. Starting from week 3 relative to the event week, the difference between treated and control firms in both Fama-French 5-factor and 3-factor cumulative abnormal returns started to stabilize between -5% and -10%. This pattern suggests that the DOJ's announcements on interlocking directorate prohibition enforcement are followed by a 5% to 10% disproportionate drop in the stock values of affected firms. Nonetheless, one noticeable pattern in figure 3 is that the drop in stock price did not start in the event week. This may be explained by the fact that the more public-facing announcement at the ABA meeting took place on Friday, and thus the market did not fully react to the news in the event week. The in-person format of the ABA meeting might have also slightly slowed down the diffusion of this news. In fact, at the daily level, the largest drop in stock returns of treated firms happened on Monday of the following week, which suggests that the stock market digested most of the information in the announcement at the ABA meeting over the weekend, and this information was immediately reflected in the stock price on the first trading day following the announcement. The fact that the treatment effect on stock returns gradually took place over a few weeks may be explained by the market's initial uncertainty about the scale and intensity of the enforcement on Section 8. Because Section 8 has not been enforced for decades, the market couldn't fully appraise the full impact of such enforcement until additional details on enforcement were gradually released in the weeks after the initial announcements.

5.3.2. Difference-in-Differences Regression on Stock Returns

Having visually shown that stock returns of treated firms underperformed control firms following the DOJ announcement, I attempt to numerically quantify the amount of loss in shareholder value by adopting the following difference-in-differences regression at the stock-weekly level:

$$R_{it} = \alpha + \beta_1 \times Treated_i \times Post_t + \delta_t + \lambda_i + \epsilon_{it} \tag{5}$$

where R_{it} includes 3 cumulative return measures. The first one is the Fama-French 5-factor cumulative abnormal return for stock i from the beginning of the pre-event window to the end of week t. The second one is the Fama-French 3-factor cumulative abnormal return for stock i from the beginning of the pre-event window to the end of week t. The third one is the raw cumulative return for stock i from the beginning of the pre-event window to the end of week t. $Treated_i$ is a dummy equal to 1 for stock

i if it were treated by the Clayton Act Section 8, as determined by the three criteria discussed earlier. $Post_t$ is a dummy equal to 1 if week t is equal to or later than the week of April 4th, 2022. I include firm and week fixed effects, and use standard errors double-clustered at the stock and week level. In the above regression, the estimate for β_1 will represent the treatment effect of the DOJ announcements on stock return. The results of this regression are presented in table 4.

The main observation from table 4 is that the treatment effect is significantly negative across all specifications. From the first column, we can see that the average treatment effect on stock returns is -7.1%, even after controlling for Fama-French 5 factors. While the second column shows a smaller treatment effect on Fama-French 3-factor cumulative abnormal returns, the statistical significance of the estimate is still at the 5% level. The third column shows that if we don't control for any asset pricing factor, the treatment effect on stock returns is -10.6%. The -7.1% drop in stock value after controlling for Fama-French five factors may seem very large at plain sight, but it is a reasonable amount if we relate it to other estimates for director values in the literature. Nguyen and Nielsen (2010) estimate the value of an independent director to be around 0.85% of the firm value by using incidences of sudden death of independent directors. Burt, Hrdlicka, and Harford (2020) show that a director contributes to firm value by 1% every year after accounting for Fama-French 5 factors. These estimates are for an average director and an average independent director, respectively. As literature (Geng et al., 2026; Barone et al., 2025; Cabezon and Hoberg, 2026) has shown numerous ways a director interlock could benefit firms, the value of an interlocking director with a product market competitor should be much higher than the 1% estimate for an average director in Burt et al. (2020).

We should notice that the 7.1% short-term drop in stock price is likely a lower bound for the cost of the Clayton Act Section 8 enforcement to shareholders for the following two reasons. First, some acts of collusion are not easily observable by outside investors, and thus the potential termination of collusion is not fully reflected in short-term stock price, even if investors know the firm is going to lose directorate interlocks. If acts of collusion are easily observable by outside investors, they should also be observed and thus already stopped by regulators, as per antitrust laws other than the Clayton Act Section 8, before the enforcement of Section 8. Therefore, even if the market knows a firm pair is going to lose an overlapping director due to the DOJ enforcement, investors will underestimate its impact on firm fundamentals until it's directly reflected in the deteriorating accounting performance. This argument suggests that the stock return of treated firms may suffer more in the long run than in the short run, due to the difficulty of observing acts of collusion through interlocking directorate. Second, although DOJ Assistant Attorney General Jonathan Kanter announced that the DOJ would start to enforce Section 8, the market might not be certain regarding the speed, scope, and intensity of the eventual enforcement. Therefore, the market response immediately after the announcement may not reflect the full impact of Section 8 enforcement on interlocking directorate among public firms. Instead, the market may gradually find out that DOJ and other regulators are very ambitious about Section 8 enforcement, especially after repeated announcements by multiple regulators and news about the actual removal of many interlocking directorates[15][16].

5.3.3. Sharp Regression Discontinuity on Stock Returns

I use a sharp regression discontinuity design to enhance the causal interpretation of the relation between DOJ announcements and stock returns. I estimate the following regression equation:

[15] https://www.justice.gov/archives/opa/pr/directors-resign-boards-five-companies-response-justice-department-concerns-about-potentially

[16] https://www.justice.gov/archives/opa/pr/justice-department-s-ongoing-section-8-enforcement-prevents-more-potentially-illegal

$$R_i = \alpha + \beta_1 \times D_i + \beta_2 \times Equity_i + \beta_3 \times D_i \times Equity_i + \epsilon_i \tag{6}$$

where D_i is a dummy equal to 1 for firm i if its latest book value of equity was greater than \$41,034,000 as of the event week of April 4th, 2022. $Equity_i$ is the latest book value of equity of firm i minus the equity threshold, in millions of dollars. The local treatment effect of interest will be captured by the estimate for β_1. The dependent variable is the daily Fama-French 5-factor abnormal return on April 11th, 2022, the trading day immediately after the DOJ announcement on April 8th, 2022. I choose to use the daily return for the next trading day because April 11th, 2022, marks the date with the greatest disparity in stock returns between the treated and control groups. Considering the Department of Justice made its announcement at the American Bar Association's spring meeting around noon, it is plausible for the stock market to respond on the subsequent day, particularly since the American Bar Association meeting is not usually an event that garners substantial market attention.

I use the same subsamples as in table 3 to run this regression. I estimate the above regression with a bias-corrected estimator and the corresponding robust standard error Calonico et al. (2014). I adopt a data-driven bandwidth selector to optimize the mean squared error (Calonico et al., 2014). Before applying the bandwidth selector, I manually drop firms with book values of equity that were either greater than 101 times the equity threshold or smaller than negative 99 times the equity threshold. I report the bias-corrected point estimates for the treated effect and the corresponding robust standard errors in table 5.

From table 5, one can observe that the estimated local treatment effect on stock returns is negative in both columns. For the subsample of firms that had at least one interlock with an above-threshold firm, the estimated treatment effect is -1.1% Fama-

French 5-factor abnormal return. Although this point estimate is much smaller than the estimated average treatment effect in table 4, the statistical significance supports the causal relation between the Section 8 enforcement and the drop in firm value. This implies that investors, at least some of them, not only efficiently process information from various sources but also can identify the equity threshold mentioned in Section 8 and factor it into their investment decisions. Given that the equity threshold is not easily noticeable to people unfamiliar with Section 8, it's surprising to observe this significant result in the regression discontinuity analysis on stock returns.

The local treatment effect on stock returns estimated in the regression discontinuity exercise is smaller than the average treatment effect estimated in the difference-in-differences analysis in table 4. There are three possible reasons for this difference. First, the regression discontinuity exercise relies on a discrete jump in stock return for firms around the threshold. Finding out the equity threshold for Section 8 is not a trivial task for investors, as it requires reading through long legal texts. It's plausible that only a small number of investors can incorporate this information on the equity threshold into their investment decisions in a short period of time, and their impact on the stock price may be limited. Second, the time horizons of stock returns are different in the difference-in-differences analysis and this regression discontinuity exercise. Stock returns considered in this regression discontinuity are only for one day after the announcement, while I consider the stock returns over two months following the events in the difference-in-differences analysis. There may be a difference in short-term and long-term stock market reactions if the market kept finding new information regarding the Section 8 enforcement, as very likely in reality. Third, regression discontinuity measures the local treatment effect, but difference-in-differences estimates the average treatment effect across all treated firms, many of which are far from the equity threshold. One may expect that firms of different sizes were differentially affected by the enforcement of Section 8.

5.3.4. Robustness Tests For Stock Performance

There may be concerns that the difference in stock returns between treated firms and control firms may be attributed to differences between industries, because treated firms are clustered in a few industries. To address this concern, I conduct a difference-in-differences analysis on stock returns, while controlling for different industries. I do so by including industry-week fixed effects in the difference-in-differences regression above. The dependent variable in this regression is the cumulative raw return since the start of the pre-event window, which starts eight weeks from the event week. I include both firm fixed effects and week-industry fixed effects in the regression. I do not include week fixed effects because the interaction between week and industry subsumes week fixed effects. Including industry-week fixed effects is equivalent to demeaning the cumulative stock returns by industry. This practice removes the impact of industry average returns on the estimated treatment effects. I also adopt two different ways to cluster the standard errors. I double-cluster the standard errors at the firm and week level and the industry and week level, respectively. To identify the industry of a firm, I use the SIC code as of the event week.

The results of this regression are presented in table 6. The first thing to note in table 6 is that the number of observations declines after including more granular industry controls because the singleton observations at the industry-week level are excluded, suggesting that there are more industries with a single firm when we define industries at a more granular scale. We can observe that after controlling for industry average returns at the 3-digit SIC level, the magnitude of the treatment effect reduces to -5.6%. This suggests that part of the treatment effect we observed in table 4 is attributed to industry-level variation in stock returns. However, the treatment effect is still negative and significant at the 1% level after controlling for industry-level differences. When I control for the industry average returns at the more granular 4-digit SIC code level, the point estimate for the treatment effect further reduces to -4.4%, and the statistical significance is at the 5% level. The results in table 6 suggest that the estimated treatment effect of DOJ announcements on firm value is not due to differential stock performance among different industries.

Another potential confounding factor behind the underperformance of treated firms is inflation news. In the morning of April 12th, 2022, the United States Bureau of Labor Statistics announced an 8.5% increase in the consumer price index year-over-year, the sharpest increase in decades, and the S&P 500 index dropped by 2.1% in this week. The week of this inflation announcement coincided with the first week after the DOJ announcements when we saw a large underperformance in the stock returns of treated firms. To investigate whether the underperformance of treated firms was due to their high exposure to inflation risk, I conduct placebo tests for three later dates in 2022 when the inflation numbers deviated from market consensus. The three placebo dates are June 10th, 2022, September 13th, 2022, and November 10th, 2022. On June 10th, 2022, the U.S. BLS reported another record-breaking 8.6% year-over-year increase in the consumer price index, and the S&P 500 index fell by 5.1% in that week. On September 13th, 2022, the consumer price index rose 8.3% year-over-year, and the S&P 500 index dropped by 4.8% in that week. Finally, on November 10th, 2022, the U.S. inflation cooled down to a 7.7% increase year-over-year. This lower-than-expected inflation number was accompanied by a 5.9% rise in the S&P 500 index in the same week. I conduct event study around these three placebo weeks using equation 4 and plot the corresponding coefficients in figure 4. The dependent variable is Fama-French 5-factor cumulative abnormal return.

From figure 4, one can see that starting from the week of June 10th, 2022 when the surprise in inflation number was in the same direction as in April 2022, treated firms outperformed control firms by almost 20%, opposite to the direction we observed after the week of April 4th, 2022. Following the weeks of September 13th, 2022 and November 10th, 2022, the treated firms and control firms had similar stock returns. The results from these three placebo tests suggest that the underperformance in the stock returns of treated firms following the announcement week was not due to inflation risks.

5.4. Channel and Heterogeneity Analysis

After showing the underperformance in stock returns for treated firms following the announcements of Section 8 enforcement, I continue to investigate the channels through which the firm could benefit from interlocking directorate with competitors. Literature has shown that product market competitors could benefit from director interlocks through coordination in the product market (Barone et al., 2025; Gopalan et al., 2026), lower R&D and investment costs (Geng et al., 2026), and market segmentation (Poberejsky, 2024). At the same time, directorate interlocks are also shown to reduce firm profitability due to lower innovation and less product differentiation and competitiveness (Cabezon and Hoberg, 2026). In this section, I inform the academic debate by doing heterogeneity analyses on both the number of interlocks and stock returns to investigate which firm characteristics are related to firms' increasing directorate interlocks and to the firm value creation by competitor interlocks, as assessed by the stock market investors.

5.4.1. Increasing Interlocks and Firm Competition

In this section, I conduct the first analysis on heterogeneity. I investigate what firm characteristics are related to increasing interlocking directorate. As shown previously in figure 1, there have been increases in both the number of interlocked firms and the number of interlocked competitors among US non-financial public firms since the early 2010s. To shed light on why firms chose to increase the number of their directorate interlocks, I plot both the number of interlocked firms and interlocked competitors again in figures 5 and 6, but this time, I separate the whole sample of firms into five groups, based on the concurrent value of total product similarities they have with all of the competitors. Different from all the analyses before, where I define competitors as firm pairs with TNIC score in the top 1%, I use the sum of TNIC scores with all competitors at TNIC-3 level, which is the top 2.05% of all firm pairs (Hoberg and Phillips, 2016). As discussed earlier, TNIC-4 level competition classification captures competition better than TNIC-3 level, but I choose to use TNIC-3 level competition classification here simply because there is a sizable portion of firms with no TNIC-4 level competitors. Therefore, if I divide firms into five groups based on total product similarity with TNIC-4 competitors, more than 20% of the firms will have zero value, resulting in unstable sorts and plots, especially when plotting the average number of interlocked firms. This problem does not exist when I use the sum of TNIC scores with TNIC-3 level competitors, because there are sufficiently many competitor pairs under this specification. However, even if I group the stocks by total similarity scores with TNIC-4 level competitors, the major patterns will be similar to those exhibited in figures 5 and 6, except for irregularities in the two subplots with the lowest similarity scores. The values on the y-axes of the two figures are computed at the end of each calendar month. Thus, I perform grouping of stocks for each month separately, based on the most up-to-date TNIC scores from the previous fiscal year.

Different from the increasing trend in the number of interlocks for the whole firm sample, as demonstrated in the top panel of figure 1, figure 5 shows that the average number of interlocked firms for the 4 groups with lower TNIC-3 total similarity remains stable between 3 and 4 throughout the years. However, the average number of interlocked firms for the top 20% of firms with the highest TNIC-3 total similarity increased by a wide margin from around 4 to 6 between 2013 and 2022. It's clear that the overall increase in the average number of interlocked firms is mainly attributed to the subset of firms that have the most homogeneous products with their competitors. Figure 6 exemplifies a much starker contrast between the top 20% of firms in total similarity and the remainder of the firms. The average number of interlocked competitors has stayed around 0 throughout the years for the bottom 80% of firms in total similarity, whereas the 20% of the firms with the highest total product similarity have experienced a six-fold increase in the number of interlocked competitors, again accounting for almost all the increase we observe in the full sample. One may argue that since both the sorting variable and the dependent variable are positively related to the number of interlocked competitors, there might be a mechanical relation that underlines the patterns in figure 6. I confirm that even when I switch the sorting variable to the average value, instead of the sum, of TNIC scores with TNIC-3 level competitors, the exact same pattern persists. Moreover, there is no straightforward mechanism where the number of interlocked firms and total similarity with competitors should be correlated, but the pattern that the top 20% of firms in total similarity accounts for the majority of the increase in the number of interlocks is also prominent in figure 5.

Figures 5 and 6 both show that the subset of firms with the most homogeneous products with competitors account for most of the increase in interlocking directorate we observe in recent years. If this increase in interlocking directorate was proactively chosen by this subset of firms, then there must be some benefit to this subset of firms by doing so. This suggests that having interlocking directorate is beneficial to the firms, and the size of this benefit is positively related to the product market closeness a firm has with its competitors.

5.4.2. Stock Performance and Firm Competition

After finding that firms with closer product market competitors contribute to most of the recent increase in interlocking directorate, I conduct a heterogeneity analysis on the magnitude of stock underperformance among treated firms. As demonstrated in the previous section, product similarity with competitors is potentially related to firms' decisions on whether to have directorate interlocks. I investigate whether the same characteristic is considered influential to firm value by stock market investors.

To perform the heterogeneity analysis on stock returns, I adopt a similar setup to that of difference-in-differences as in section 5.3.2, except that I add a triple interaction term with the variables on the firm characteristics of interest. The regression equation is as follows:

$$\begin{aligned} R_{it} = \alpha + \beta_1 \ \times \ Treated_i \times Post_t \\ + \beta_2 \times Treated_i \times Post_t \times Competition_i + \delta_t + \lambda_i + \epsilon_{it} \end{aligned} \tag{7}$$

where R_{it} includes 3 cumulative return measures: including the Fama-French 5-factor cumulative abnormal return, the Fama-French 3-factor cumulative abnormal return, and the raw cumulative return for stock i from the beginning of the pre-event window to the end of week t. $Treated_i$ is a dummy variable equal to 1 for stock i if it was treated by the Clayton Act Section 8, as determined by the three criteria discussed in section 5.1. $Post_t$ is a dummy variable equal to 1 if week t is equal to or later than the week of April 4th, 2022. $Competition_i$ is one of the two variables: the sum of TNIC scores with all TNIC-4 level competitors or the number of TNIC-4 level competitors. The competition variables are from fiscal year 2021, which were up to date as of April 2022. I include firm and week fixed effects and use standard errors double-clustered at the stock and week level. The competition variables are not included as a standalone term in the regression equation because they would have been absorbed by the firm fixed effect. The results of this regression are presented in table 7.

From table 7, we can observe that the triple interaction terms with the two competition variables have negative coefficients in all six columns for all three measures of stock returns. The negative coefficients for the triple interaction with the sum of product similarity scores with all TNIC-4 competitors mean that, among the treated firms, the underperformance of the stock returns following the announcement events is negatively related to the total product similarity with competitors. It shows that the treated firms with more homogeneous products with competitors had a lower stock return than other treated firms following the shock. The coefficients for the triple interaction term with the number of competitors are also negative across the three return specifications, meaning that a treated firm had a lower stock return after the announcements if it operated in product markets with more competitors. Importantly, these negative coefficients of triple interaction terms should be interpreted as the heterogeneity in treatment effects among treated firms, instead of the difference between treated firms and control firms. Moreover, we observe that after including the triple interaction terms in the regressions, the coefficients of the double interaction terms become positive, opposite to the negative signs in table 4. This suggests that the negative treatment effect we observe in table 4 is fully explained by product market competition.

There may be concerns that the treatment classification could be correlated with the two competition variables because the total similarity is potentially related to the interlocking directorate among competitors, as shown in figure 6, making estimations in table 7 unreliable. To address this concern, I perform cross-sectional regressions on post-event stock returns with the subset of treated firms only. I run the following cross-sectional regression:

$$R_i = \alpha + \beta_1 \times Competition_i + \beta_2 \times Controls_i + \epsilon_i \tag{8}$$

where R_i is the Fama-French 5-factor cumulative abnormal return for stock i from the end of treatment week to the end of the third week after the treatment. $Competition_i$ is one of the two variables on the product market competition faced by firm i. The two competition variables are the same as defined in the difference-in-differences regression in table 4. Control variables include market capitalization, book-to-market ratio, debt-to-equity ratio, volatility, and Amihud liquidity ratio. Market capitalization, book-to-market ratio, and debt-to-equity ratio are the latest as of the beginning of the treatment week. Volatility and Amihud illiquidity ratio are computed from stock returns and trading volumes in the 21-trading day period immediately before the treatment week. All firms included in this regression are treated firms, as defined in section 5.1.

I choose to compute the cross-sectional return from the end of the treatment week to the end of the third week after treatment because this is the period when the treatment effect on stock returns is the most prominent, as shown in figure 3. Figure 3 shows that the difference in stock returns between the treatment and control groups stabilized after the end of week 3. At the daily level, the largest movement in stock price took place on the first trading day of week 1, which is also included in the sample period for this cross-sectional regression. This large movement on the first trading day following the ABA meeting announcement fits well with the policy shock I propose in this paper, because the ABA meeting on Friday was more public-facing than the Spring Enforcer meeting on Monday, which targeted regulators around the globe. It's reasonable for the market to digest the policy shock and identify the affected firms over the weekend and then react on the following trading day. Therefore, the sample period from the end of the treatment week to the end of the third week after treatment is likely to capture the majority of the treatment effect on the stock price.

The results from the cross-sectional regression above are presented in table 8. The first observation from table 8 is that both the total similarity with competitors and the number of TNIC-4 level competitors are negatively correlated with the post-event stock returns of treated firms. This is consistent with the observation in difference-in-differences regressions in table 7, suggesting that the results in table 7 are not driven by a potential correlation between the treated dummy and competition variables. Moreover, the inclusion of control variables does not affect the estimates for the main competition variables. It's not surprising that the coefficients for the control variables are not significant because most of the factors affecting stock returns are already controlled for in the computation of Fama-French 5-factor abnormal returns. The results from this cross-sectional regression corroborate with the results from the previous difference-in-differences regression and confirm the finding that among treated firms, those with more homogeneous competing products and more competitors had lower stock returns following the announcements on Section 8 enforcement.

Overall, there are two major findings from the above analysis. First, firms that had the most homogeneous products with competitors were the main driving force behind the rapidly increasing interlocking directorate among US non-financial public firms in the recent decade. Second, among the firms affected by the Section 8 enforcement, those with more homogeneous competing products and more competitors were the ones experiencing larger drops in firm value. These results are consistent with the theory that competing firms collude in the product market through sharing common directors. As suggested by canonical economic models on competition, it's more beneficial for firms to collude if they operate in more similar markets. However, coordinating and sustaining such collusion requires a communication channel and a monitoring channel. Without a covert communication channel, it might be hard for competitors to coordinate on collusion without attracting attention from regulators. Without a reliable monitoring channel, it's hard to detect deviations, and thus sustaining collusion could depend on many restrictive conditions (Harrington and Skrzypacz, 2011). The interlocking directorate could serve as both a private communication channel and an effective monitoring channel because directors could both privately interact with management teams and reliably monitor firm strategies and product market actions (Antón, Ederer, Giné, and Schmalz, 2023; Barone et al., 2025). Hence, the product market collusion channel could explain the findings in this section that the coexistence of common directors and product homogeneity brings value to the firm. This suggests that an increasingly important way a director could provide value to the firm is by acting as a channel for covert two-way information exchange between rival firms.

Moreover, the findings in the above heterogeneity analysis are inconsistent with the alternative theory of lower board quality. As the enforcement of Section 8 reduced the number of firms a director can simultaneously work for, the total supply of qualified directors would decrease as a result. Since, in equilibrium, a firm will hire the best quality directors who can provide the highest monitoring and advisory value, the lower director supply will result in many firms replacing current high-quality directors with lower-quality ones. However, this expected reduction in monitoring and advisory value should not vary with the level of product market competition faced by each firm, contradicting the findings in the heterogeneity analysis. To further test the alternative hypothesis of board quality, I conduct an analysis of the relationship between stock returns around the announcement event and the quality of overlapping directors in the next section.

5.4.3. Stock Returns and Board Quality

In this section, I directly test the alternative hypothesis that firms experience negative returns after the announcement because they are expected to lose high quality directors as a result of the regulation enforcement. If this hypothesis is true, then we should expect firms with higher-quality overlapping directors to lose more value, assuming the new directors who replace them are of average quality. Even if this assumption is violated in reality, our expectation for post-event stock return still holds true as long as the market held the same assumption at the time of the announcement.

To test the relationship between post-event stock returns and the quality of the overlapping director, I adopt a similar triple-interaction difference-in-differences setup as in section 5.4.2. The regression specification is as follows:

$$\begin{aligned} R_{it} = \alpha + \beta_1 \times Treated_i \times Post_t \\ + \beta_2 \times Treated_i \times Post_t \times BoardQuality_i + \delta_t + \lambda_i + \epsilon_{it} \end{aligned} \tag{9}$$

where R_{it} is the Fama-French 5-factor abnormal returns since the beginning of the pre-event window. $BoardQuality_i$ is the firm-level proxy for the quality of its directors who simultaneously sit on the board of an above-threshold TNIC-4 level product market competitor as of the end of March 2022. As the literature has suggested that director age, tenure length at the firm, education level, and the size of professional connections are correlated with the firm value enhancement by the director (Huang and Hilary, 2018; Masulis, Wang, Xie, and Zhang, 2025; Gao and Huang, 2024; Omer, Shelley, and Tice, 2020; Intintoli, Kahle, and Zhao, 2018), I use these measures as proxies for director quality. Age is the difference between the year 2022 and the director's year of birth. Director tenure length is defined as the number of calendar days since this director first served at this firm until the event date. This could capture the value derived from this director's familiarity with the firm operations. Education level captures the highest level of education attained by the director, and it is defined as 1 for undergraduate education, 2 for master's level degrees, 3 for doctoral degrees, and 0 otherwise. Network size is the size of the professional connections of the director. All these measures are obtained from director profiles in the BoardEx dataset. For this regression, I aggregate these director level or director-firm level quality measures to the firm level as discussed below.

Because Section 8 enforcement only increased the replacement risk for the directors who simultaneously sit on the boards of product market competitor firms that had a book value of equity above the threshold, I compute the firm-level measure for director quality by summing the quality proxies for all its directors who simultaneously serve as directors for at least one above-threshold product market competitor. If the director quality hypothesis is true, the coefficients for these triple interaction terms should be significant because the treated firms with higher-quality overlapping directors are expected to lose more value following the regulation enforcement. The standard errors are double clustered at the firm and week levels.

The results of the regressions are reported in table 9. As demonstrated in table 9, the triple interaction terms in all four columns are not significant, whereas the double interaction terms remain negative and are significant in most columns. These results suggest that among firms treated by Section 8 enforcement, the quality of the overlapping directors is not correlated with the post-event stock returns. Therefore, the potential loss of high-quality directors as a result of the regulation enforcement does not explain the post-event drop in the stock prices of the treated firms. Overall, the test results in the heterogeneity analyses are not consistent with the hypothesis that treated firms lose value because they could lose the advisory or monitoring value brought by high-quality directors following the shock. In unreported regressions, I also use the average overlapping director quality measures, instead of the sum of quality proxies, and the results are qualitatively similar.

6. Conclusion

In this paper, I propose a novel policy shock to the enforcement probability regarding board interlocks between product market competitors. In the first part of the paper, I validate this shock by confirming its negative impact on the number of interlocks among U.S. non-financial public firms through both an event study and a regression discontinuity exercise. Then I evaluate the stock market reaction to this policy shock and confirm that this policy shock negatively affected the value of the treated firms. Lastly, I find that the recent rise in interlocking directorate is attributed to the firms facing the most intense competition in the product market and that the value of such interlocks is higher for firms facing higher levels of competition. These findings support the theory that competing firms collude in the product market by sharing directors.

I make two major contributions to the literature. First, I add to the corporate governance literature by showing that an important function of the corporate board could be enabling covert two-way information flow between firms, especially those competing in the product market. Second, I provide insight to the contemporaneous literature on interlocking directorate between competitors by measuring the overall value of such interlocks to the firms through stock prices.

Appendix A. Tables

Table 1: Summary Statistics for Treated and Control Group in Event Study

This table presents the summary statistics on basic firm characteristics and interlocking directorate measures for the firms in my sample. The sample is divided into two groups: treated group and control group. Treated group and control group are defined as in section 5.1. The beginning number of interlocking directorates are cross-sectional as of March 31st, 2022. The change in the number of interlocking directorate measures the change from March 31st, 2022 to December 31st, 2022. The last column presents the summary statistics for all firms in my whole sample. Mean values and standard deviations (in parentheses) are reported, except in the first row.

	Control	Treated	Total
N	2,751 (89.2%)	334 (10.8%)	3,085 (100.0%)
Book value of equity (in million dollars)	3639.64 (15946.17)	532.12 (1368.22)	3299.00 (15084.61)
Debt-to-equity ratio	1.22 (13.85)	0.62 (1.05)	1.16 (13.08)
Book-to-market ratio	0.49 (0.77)	0.48 (0.42)	0.49 (0.74)
Beginning number of directors	9.49 (2.79)	9.15 (1.74)	9.46 (2.70)
Beginning number of common directors	2.97 (2.41)	4.93 (1.94)	3.19 (2.44)
Beginning number of interlocked firms	3.88 (3.48)	7.87 (3.89)	4.32 (3.73)
Beginning number of interlocked above-threshold firms	3.43 (3.28)	6.75 (3.75)	3.79 (3.49)
Beginning number of interlocked above-threshold competitors	0.07 (0.56)	4.09 (3.15)	0.50 (1.71)
Change in number of directors	0.01 (1.22)	0.00 (0.96)	0.01 (1.20)
Change in number of common directors	-0.07 (0.89)	-0.13 (0.99)	-0.08 (0.90)
Change in number of interlocked firms	-0.16 (1.28)	-0.68 (1.60)	-0.22 (1.33)
Change in number of interlocked above-threshold firms	-0.23 (1.23)	-0.92 (1.68)	-0.31 (1.31)
Change in number of interlocked above-threshold competitors	-0.01 (0.29)	-0.62 (1.31)	-0.08 (0.55)

Table 2: Monthly Difference-in-Differences on Interlocking Directorate

This table shows the change in the interlocking directorate among US public firms around April 2022, when Department of Justice made a series of announcements on stepping up the enforcement of the Clayton Act Section 8. The regression is at firm-monthly level. The sample period for this regression is from October 2021 (6 months before event month) to October 2022 (6 months after event month). The firm sample includes all non-financial public firms with non-missing equity observations from Compustat. The dependent variables reported in the table include: (1) number of linked firms through interlocking directorate, (2) number of interlocked firms that had the latest book value of equity greater than the equity threshold, (3) number of interlocked competitors, as defined by TNIC-4 similarity measure, that had the latest book value of equity greater than the equity threshold. These variables are computed for the end of each calendar month. Post dummy is equal to 1 if the month is later than or equal to April 2022, and equal to 0 if earlier than April 2022. Treated dummy is equal to 1 if this firm is involved in any firm pair that simultaneously satisfies the following 3 conditions: (1) both firms have equity value above threshold; (2) two firms have at least 1 common director; (3) two firms are identified as product market competitors by TNIC-4 measure. Firm and month fixed effects are included. Standard errors are double-clustered at firm and month level. *, **, *** denote 10%, 5%, 1% significance respectively.

Dependent Variable:	(1)	(2)	(3)
Post dummy × Treated dummy	-0.203**	-0.377***	-0.352***
	(0.085)	(0.105)	(0.080)
Constant	4.227***	3.708***	0.513***
	(0.005)	(0.006)	(0.005)
Firm FE	YES	YES	YES
Month FE	YES	YES	YES
R-squared	0.96	0.96	0.98
N	41,730	41,730	41,730

Table 3: Regression Discontinuity on Change in Interlocking Directorate

This table presents results from sharp regression discontinuity analysis on the change in interlocking directorate. Two subsamples of firms are used. First subsample is the firms with at least one interlocked firm that was above the equity threshold, as of March 31st, 2022. Second subsample is the firms with at least one interlocked product market competitor, as identified by TNIC-4 measure, that was above the equity threshold, as of March 31st, 2022. First dependent variable is the change in the number of interlocked firms that was above the equity threshold. Second dependent variable is the change in the number of interlocked competitors that was above the equity threshold. Changes are measured in two horizons. First is the change from March 31st, 2022 to April 30th, 2022. Second is the change from March 31st, 2022 to May 31st, 2022. Treatment effect is estimated with bias correction and the corresponding robust standard errors (Calonico et al., 2014). Mean squared error-optimal bandwidth selection (Calonico et al., 2014) is applied after manually excluding firms with equity greater than 101 times the equity threshold or less than negative 99 times the equity threshold. *, **, *** denote 10%, 5%, 1% significance respectively.

Subsample:	Linked Above-Threshold Firm > 0		Linked Above-Threshold Competitor > 0	
Dependent variable:	Δ Linked Above-Threshold Firm		Δ Linked Above-Threshold Competitor	
Horizon:	April	May	April	May
Treatment effect	-0.056	-0.149**	-0.108*	-0.273*
	(0.043)	(0.061)	(0.059)	(0.163)
N	1,273	1,382	251	248

Table 4: Underperformance of Stock Returns for Treated Firms After DOJ Announcements

This table quantifies the difference in stock returns after the DOJ announcements between firms that were treated by the Clayton Act Section 8 and firms that were not treated. I adopt a difference-in-differences design. The regression is at firm-week level. The sample period is from 8 weeks before event week to 8 weeks after event week. Event week is the week of 4th April, 2022. The dependent variables are Fama-French 5-factor cumulative abnormal return, Fama-French 3-factor cumulative abnormal return, and cumulative raw return. All cumulative returns are computed from the beginning of week -8 to the end of the week of observation. Fama-French factor loadings are computed for each stock using all daily observations in 2021, the year before DOJ announcements. Treated dummy is equal to 1 if the stock was treated, as defined in section 5.1, by the Clayton Act Section 8. Post dummy is equal to 1 if the observation week is equal to or after the week of April 4th, 2022, and equal to 0 otherwise. Firm and week fixed effects are included. Standard errors are double clustered at firm and week level. *, **, *** denote 10%, 5%, 1% significance respectively.

Dependent Variable:	FF5 CAR	FF3 CAR	Cumulative Return
Treated dummy $\times$ Post dummy	-0.071***	-0.053**	-0.106***
	(0.022)	(0.020)	(0.023)
Constant	-0.027***	0.004***	-0.038***
	(0.001)	(0.001)	(0.001)
Firm FE	YES	YES	YES
Week FE	YES	YES	YES
R-squared	0.75	0.73	0.71
N	51,725	51,725	52,392

Table 5: Regression Discontinuity on Stock Returns

This table presents results from sharp regression discontinuity analysis on the stock return. Two subsample of firms are used: (1) firms with at least one interlocked firm that was above the equity threshold, as of March 31st, 2022; (2) firms with at least one interlocked product market competitor, as identified by TNIC-4 measure, that was above the equity threshold, as of March 31st, 2022. The dependent variable is Fama-French 5-factor abnormal return on April 11th, 2022, the trading day immediately following the DOJ announcement on April 8th, 2022. Treatment effect is estimated with bias correction and the corresponding robust standard errors (Calonico et al., 2014). Mean squared error-optimal bandwidth selection (Calonico et al., 2014) is applied after manually excluding firms with equity greater than 101 times the equity threshold or less than negative 99 times the equity threshold. *, **, *** denote 10%, 5%, 1% significance respectively.

Dependent variable:	Fama-French 5-Factor Abnormal Return	
Subsample:	(1)	(2)
Treatment effect	-0.011*	-0.005
	(0.006)	(0.014)
N	1,551	272

Table 6: Underperformance of Stock Returns After Controlling For Industry

This table quantifies the difference in stock returns after the DOJ announcements between treated and control firms, after controlling for average returns in different industries. I adopt a difference-in-differences design. The regression is at firm-week level. The sample period is from 8 weeks before event week to 8 weeks after event week. Event week is the week of 4th April, 2022. The dependent variable is cumulative raw return from the beginning of week -8 to the end of the week of observation. Treated dummy is equal to 1 if the stock was treated, as defined in section 5.1, by the Clayton Act Section 8. Post dummy is equal to 1 if the observation week is equal to or after the week of April 4th, 2022, and equal to 0 otherwise. Both firm and industry-week fixed effects are included. The industry of a firm is classified by its 3-digit or 4-digit SIC codes as of the event week. Standard errors are either double clustered at firm and week level or double clustered at industry and week level. *, **, *** denote 10%, 5%, 1% significance respectively.

Dependent Variable:	Cumulative Return			
Industry classification:	SIC-3		SIC-4	
Treated dummy × Post dummy	-0.056***	-0.056***	-0.044**	-0.044**
	(0.018)	(0.018)	(0.018)	(0.019)
Constant	-0.042***	-0.042***	-0.046***	-0.046***
	(0.001)	(0.001)	(0.001)	(0.001)
Firm FE	YES	YES	YES	YES
Week-SIC FE	YES	YES	YES	YES
Clustering	Firm-Week	SIC-Week	Firm-Week	SIC-Week
R-squared	0.76	0.76	0.76	0.76
N	50,898	50,898	47,818	47,818

Table 7: Product Market Competition and Stock Returns After DOJ Announcements

This table demonstrates the relation between product market competition and stock underperformance following the announcement events. I adopt a difference-in-differences design. The regression is at firm-week level. The sample period is from 8 weeks before event week to 8 weeks after event week. Event week is the week of 4th April, 2022. The dependent variables are Fama-French 5-factor cumulative abnormal return, Fama-French 3-factor cumulative abnormal return, and cumulative raw return. All cumulative returns are computed from the beginning of week -8 to the end of the week of observation. Fama-French factor loadings are computed for each stock using all daily observations in 2021, the year before DOJ announcements. Treated dummy is equal to 1 if the stock was treated, as defined in section 5.1, by the Clayton Act Section 8. Post dummy is equal to 1 if the observation week is equal to or after the week of April 4th, 2022, and equal to 0 otherwise. Firm-level competitor total similarity is the sum of TNIC scores for each firm with all of its TNIC-4 level competitors in fiscal year 2021. Firm-level number of competitors is the number of its TNIC-4 level competitors in fiscal year 2021. Firm and week fixed effects are included. Standard errors are double clustered at firm and week level. *, **, *** denote 10%, 5%, 1% significance respectively.

Dependent Variable:	FF5 CAR		FF3 CAR		Cumulative Return	
Treated × Post	0.080**	0.084**	0.086***	0.089***	0.059**	0.064**
	(0.030)	(0.031)	(0.027)	(0.027)	(0.026)	(0.027)
Treated × Post × Competitor Total Similarity	-0.002***		-0.002***		-0.002***	
	(0.000)		(0.000)		(0.000)	
Treated × Post × Number of Competitors		-0.001***		-0.000***		-0.001***
		(0.000)		(0.000)		(0.000)
Constant	-0.027***	-0.027***	0.004***	0.004***	-0.038***	-0.038***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Firm FE	YES	YES	YES	YES	YES	YES
Week FE	YES	YES	YES	YES	YES	YES
R-squared	0.75	0.75	0.73	0.73	0.71	0.71
N	51,725	51,725	51,725	51,725	52,392	52,392

Table 8: Return Heterogeneity Among Treated Firms

This table presents results from cross-sectional regressions of stock returns after DOJ announcements for the subsample of firms treated by the announcements. The dependent variable is the Fama-French 5-factor cumulative abnormal return from the end of the treatment week to the end of the third week after the treatment week. Fama-French factor loadings are computed for each stock using all daily observations in 2021, the year before DOJ announcements. The main independent variables include the sum of product similarity scores with all TNIC-4 level competitors and the number of TNIC-4 level competitors. TNIC scores are from fiscal year 2021. The control variables include log of firm market capitalization, book-to-market ratio, debt-to-equity ratio, volatility, and Amihud illiquidity measure (Amihud, 2002). Both volatility and Amihud illiquidity measure are computed using daily data from the period of 21 trading days immediately before treatment week. All other control variables are from the latest data as of the beginning of treatment week. All firms included in this regression are treated by DOJ announcements, as defined in section 5.1. *, **, *** denote 10%, 5%, 1% significance respectively.

Dependent Variable:	FF5 CAR in weeks 1 to 3			
Competitor Total Similarity	-0.001***	-0.001***		
	(0.000)	(0.000)		
Number of Competitors			-0.000***	-0.000***
			(0.000)	(0.000)
Log(Market Cap)		0.003		0.002
		(0.011)		(0.011)
Book-to-Market		0.046		0.044
		(0.030)		(0.030)
Debt-to-Equity		-0.011		-0.011
		(0.011)		(0.011)
Amihid Illiquidity		-0.000		-0.000
		(0.000)		(0.000)
Volatility		-0.052		-0.056
		(0.497)		(0.497)
Constant	-0.004	-0.035	-0.001	-0.031
	(0.023)	(0.100)	(0.023)	(0.101)
R-squared	0.07	0.09	0.07	0.08
N	334	334	334	334

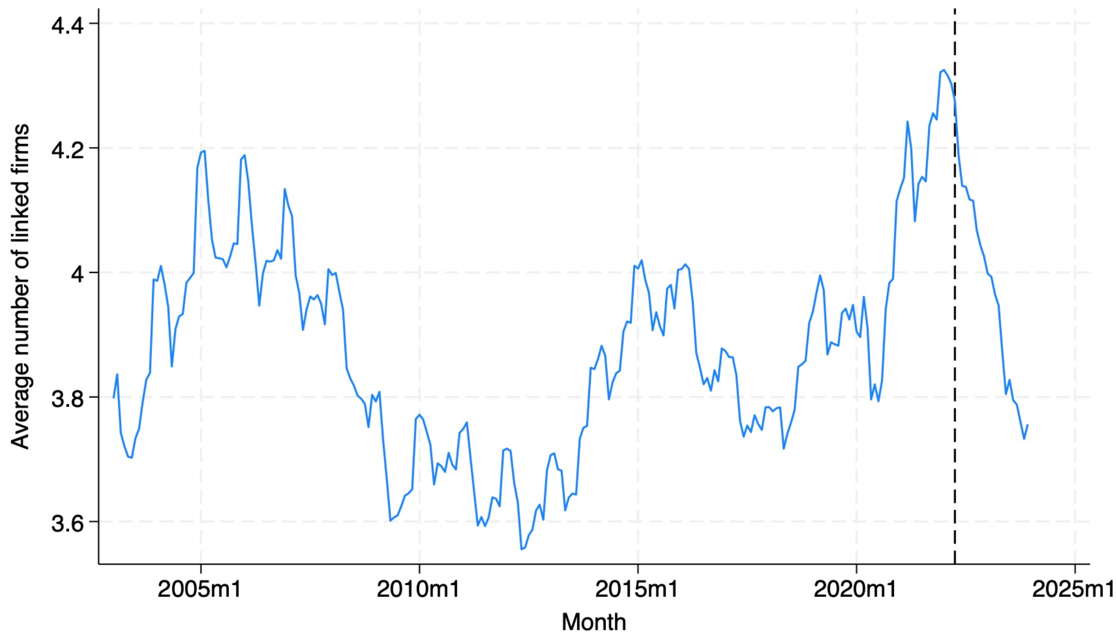
Table 9: Stock Returns and Director Quality

This table demonstrates the relation between stock returns around the announcement event and the quality of the common directors. The regression is at firm-week level. The sample period is from 8 weeks before event week to 8 weeks after event week. Event week is the week of 4th April, 2022. The dependent variable is Fama-French 5-factor cumulative abnormal return. The cumulative return is computed from the beginning of week -8 to the end of the week of observation. Fama-French factor loadings are computed for each stock using all daily observations in 2021, the year before DOJ announcements. Treated dummy is equal to 1 if the stock was treated, as defined in section 5.1, by the Clayton Act Section 8. Post dummy is equal to 1 if the observation week is equal to or after the week of April 4th, 2022, and equal to 0 otherwise. The additional interaction terms are proxies for the quality of the directors who simultaneously sit on the board of an above-threshold TNIC-4 level product market competitor. The quality proxies include the sums of ages, tenure lengths, education levels, and network sizes of the overlapped directors. Firm and week fixed effects are included. Standard errors are double clustered at firm and week level. *, **, *** denote 10%, 5%, 1% significance respectively.

Dependent Variable:	FF5 CAR			
Treated × Post	-0.070*	-0.079***	-0.056*	-0.046
	(0.033)	(0.023)	(0.031)	(0.028)
Treated × Post × Sum of ages	-0.000			
	(0.000)			
Treated × Post × Sum of tenure lengths		0.000		
		(0.000)		
Treated × Post × Sum of education levels			-0.002	
			(0.003)	
Treated × Post × Sum of network sizes				-0.000
				(0.000)
Constant	-0.027***	-0.027***	-0.027***	-0.027***
	(0.001)	(0.001)	(0.001)	(0.001)
Firm FE	YES	YES	YES	YES
Week FE	YES	YES	YES	YES
R-squared	0.75	0.75	0.75	0.75
N	51,725	51,725	51,725	51,725

Appendix B. Figures

Fig. 1. Monthly Evolution of Interlocking Directorate Among US Public Firms
This figure shows the monthly time-series plots on the interlocking directorate among all US non-financial public firms. The first panel shows the average number of interlocked non-financial firms for each US non-financial public firm at the end of each month. The second panel shows the average number of interlocked non-financial competitors for each US non-financial public firm at the end of each calendar month. Competitors are defined at TNIC-4 level. The vertical dashed line denotes April 2022, when 2 DOJ announcements on the Clayton Act Section 8 enforcement took place.



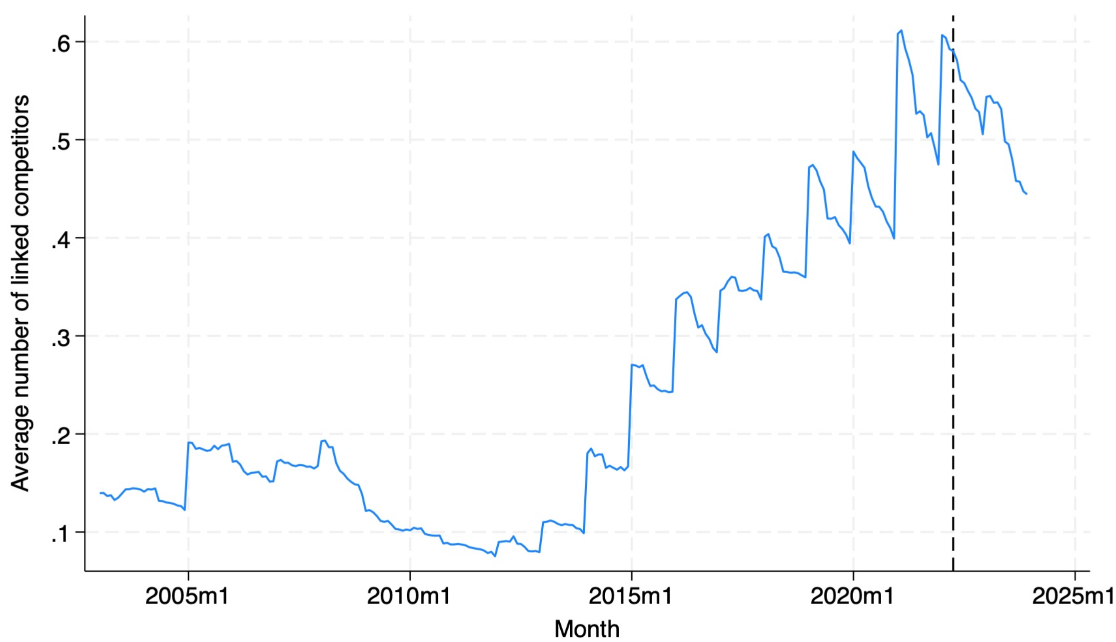


Fig. 2. Monthly Event Study Plot for Interlocking Directorate Around Announcements

This event study plot draws the evolution of differences in interlocking directorate connections between stocks treated by the Clayton Act Section 8 and stocks not treated. The event study regression is at firm-month level. The sample period for this regression is from October 2021 (6 months before event month) to October 2022 (6 months after event month). The firm sample includes all non-financial public firms with non-missing book value of equity observations from Compustat. Treatment classification is derived from 3 criteria: book value of equity, interlocking directorate, and product market competition, as described in section 5.1. The x-axis is the month relative to the month of event, where April 2022 is labeled as 0. The first panel shows the evolution of differences between treated and control firms in the number of firms interlocked through common directors. The second panel shows the evolution of differences between treated and control firms in the number of above-threshold firms linked through common directors. The third panel shows the evolution of differences between treated and control firms in the number of interlocked above-threshold product market competitors, as identified by TNIC-4 measure. The vertical dashed line denotes the month of event, April 2022. Standard errors are double-clustered at firm and month level. The confidence intervals are at 95% level.

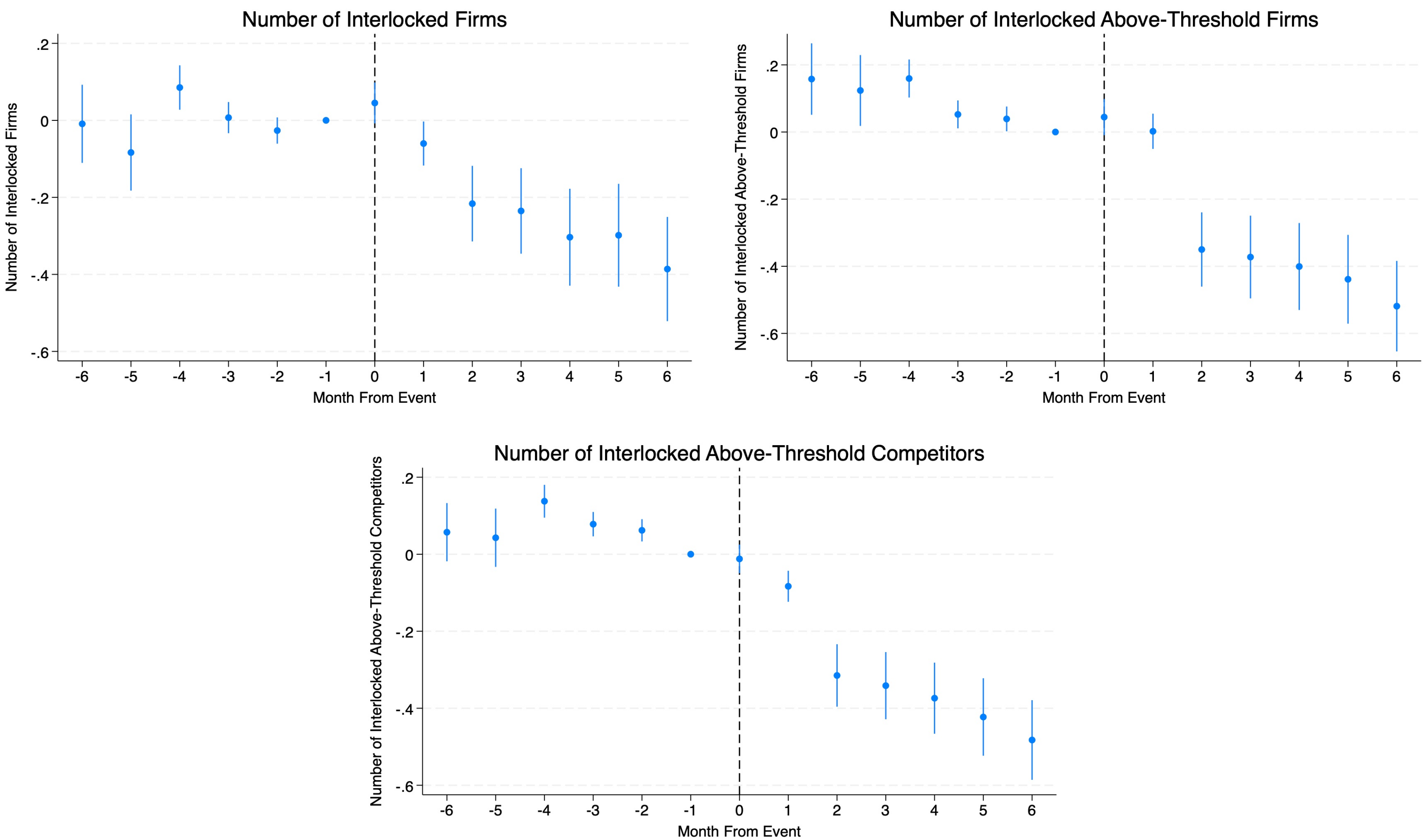


Fig. 3. Weekly Event Study Plot for Stock Return Around DOJ Announcements
This event study plot draws the evolution of differences in Fama-French 5-factor and 3-factor cumulative abnormal returns between stocks treated by the Clayton Act Section 8 and stocks not treated. The regression is at firm-week level. The sample period is from 8 weeks before event week to 8 weeks after event week. The firm sample includes all non-financial public firms with latest non-missing book value of equity observations from Compustat. Treatment classification is derived from 3 criteria: book value of equity, interlocking directorate, and product market competition. The y-axis is the Fama-French cumulative abnormal return since the beginning of the pre-window in decimal. The x-axis is the relative week from the beginning of the window. The vertical dashed line denotes the event week, the week of April 4th, 2022. The coefficients plotted represent the difference in FF5 and FF3 CAR between treated and control groups in corresponding weeks. The confidence intervals are at 95% level. Standard errors are double clustered at firm and week level.

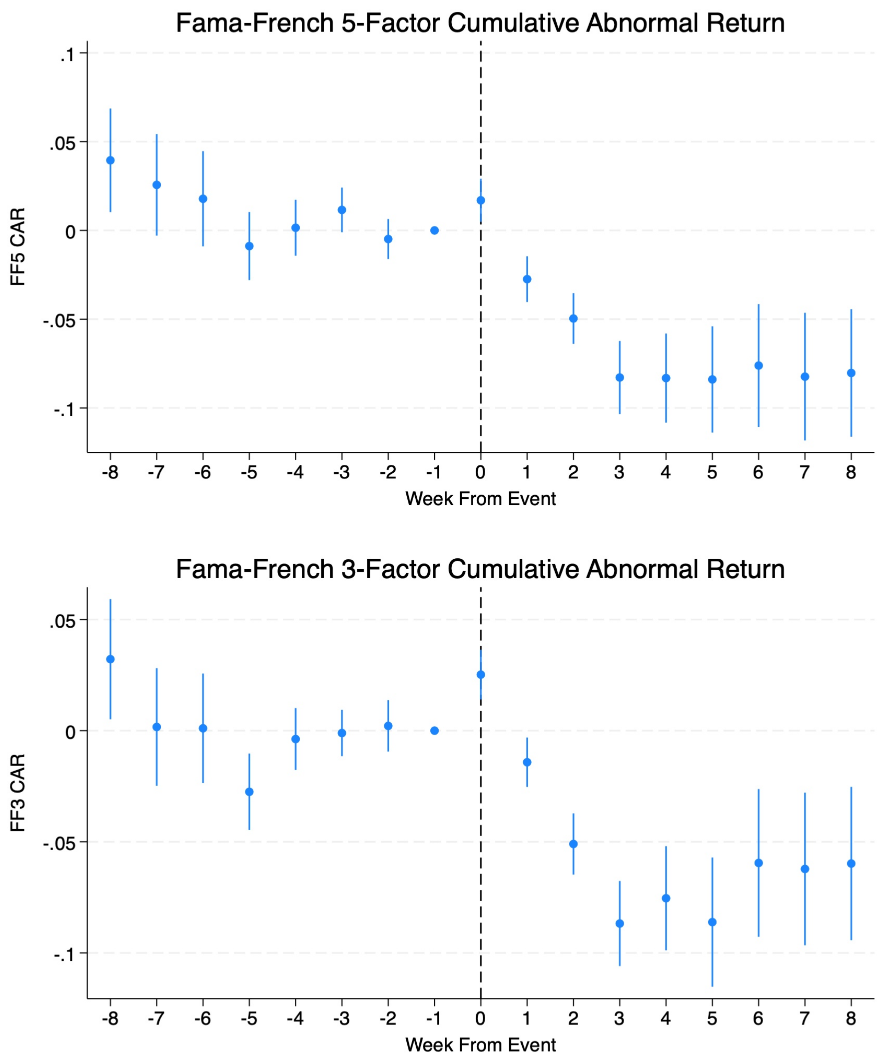


Fig. 4. Weekly Event Study Plot for Stock Return Around Inflation News Events

This event study plot draws the evolution of differences in Fama-French 5-factor cumulative abnormal returns between stocks treated by the Clayton Act Section 8 and stocks not treated around three inflation news events in 2022. The regression is at firm-week level. The sample period is from 8 weeks before event week to 8 weeks after event week. The firm sample includes all non-financial public firms with latest non-missing book value of equity observations from Compustat. Treatment classification is derived from 3 criteria: book value of equity, interlocking directorate, and product market competition. The y-axis is the Fama-French cumulative abnormal return since the beginning of the pre-window in decimal. The x-axis is the relative week from the beginning of the window. The weeks of three inflation report events are used. The dates of inflation news events are June 10th, 2022, September 13th, 2022, and November 10th, 2022. The vertical dashed lines denote the event weeks. The coefficients plotted represent the difference in FF5 CAR between treated and control groups in corresponding weeks. The confidence intervals are at 95% level. Standard errors are double clustered at firm and week level.

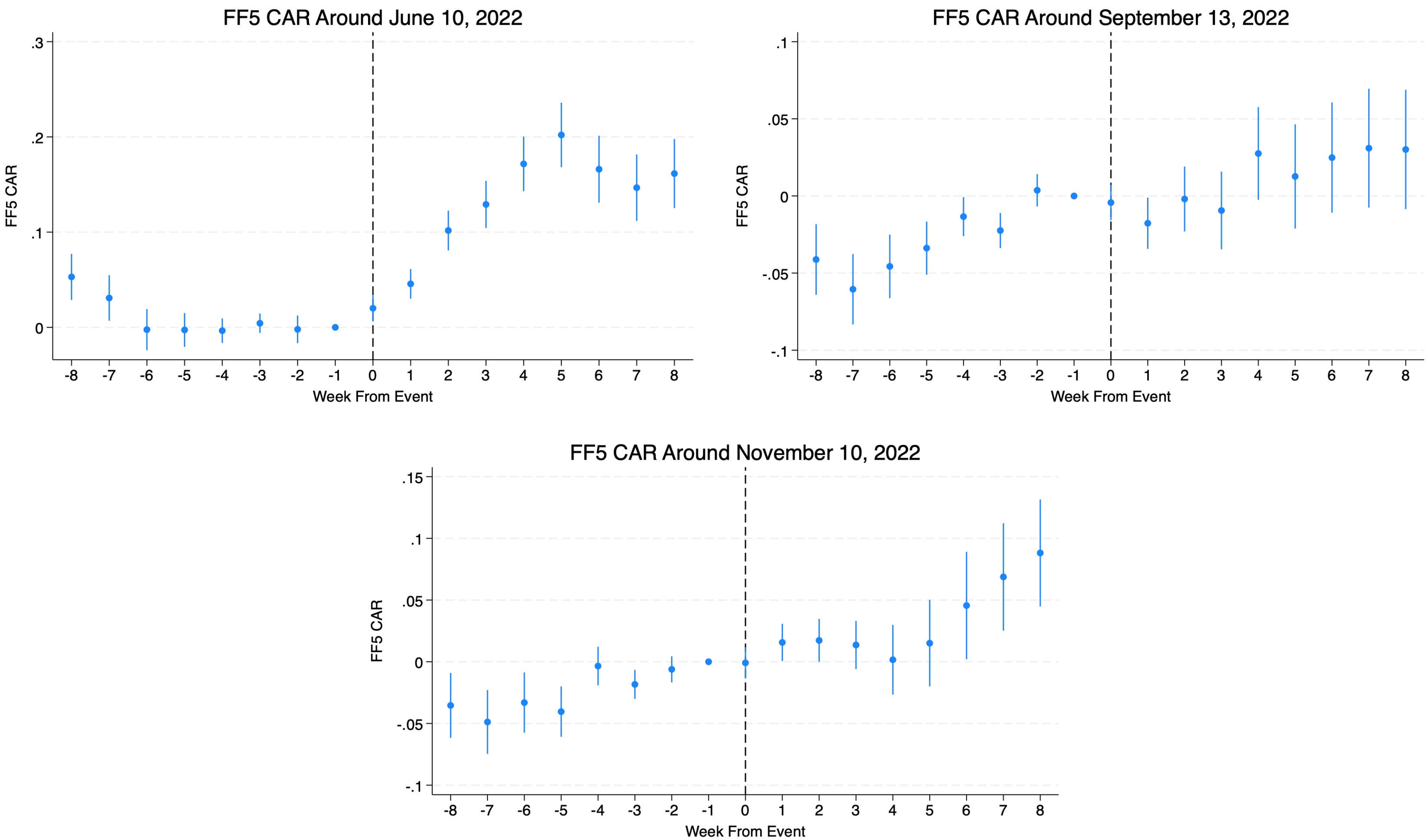


Fig. 5. Evolution of Interlocked Firms Grouped by TNIC-3 Total Similarity

This plot shows different trends in the number of interlocked firms among groups of stocks with different product similarity with competitors. For each month, I separate all non-financial US public firms into 5 groups based on the firm-month level sum of product similarity scores between the firm and all of its TNIC-3 competitors. Then for each group in each month, I compute the average number of interlocked non-financial firms across all firms in the group. The y-axis in the plot is the average number of interlocked firms. The x-axis is the calendar month. The number above each subplot represents the corresponding quintile for each subplot, with 1 denoting the group with the lowest product similarity with competitors and 5 denoting the group with the highest product similarity. Vertical dashed line denotes April 2022, the month of event.

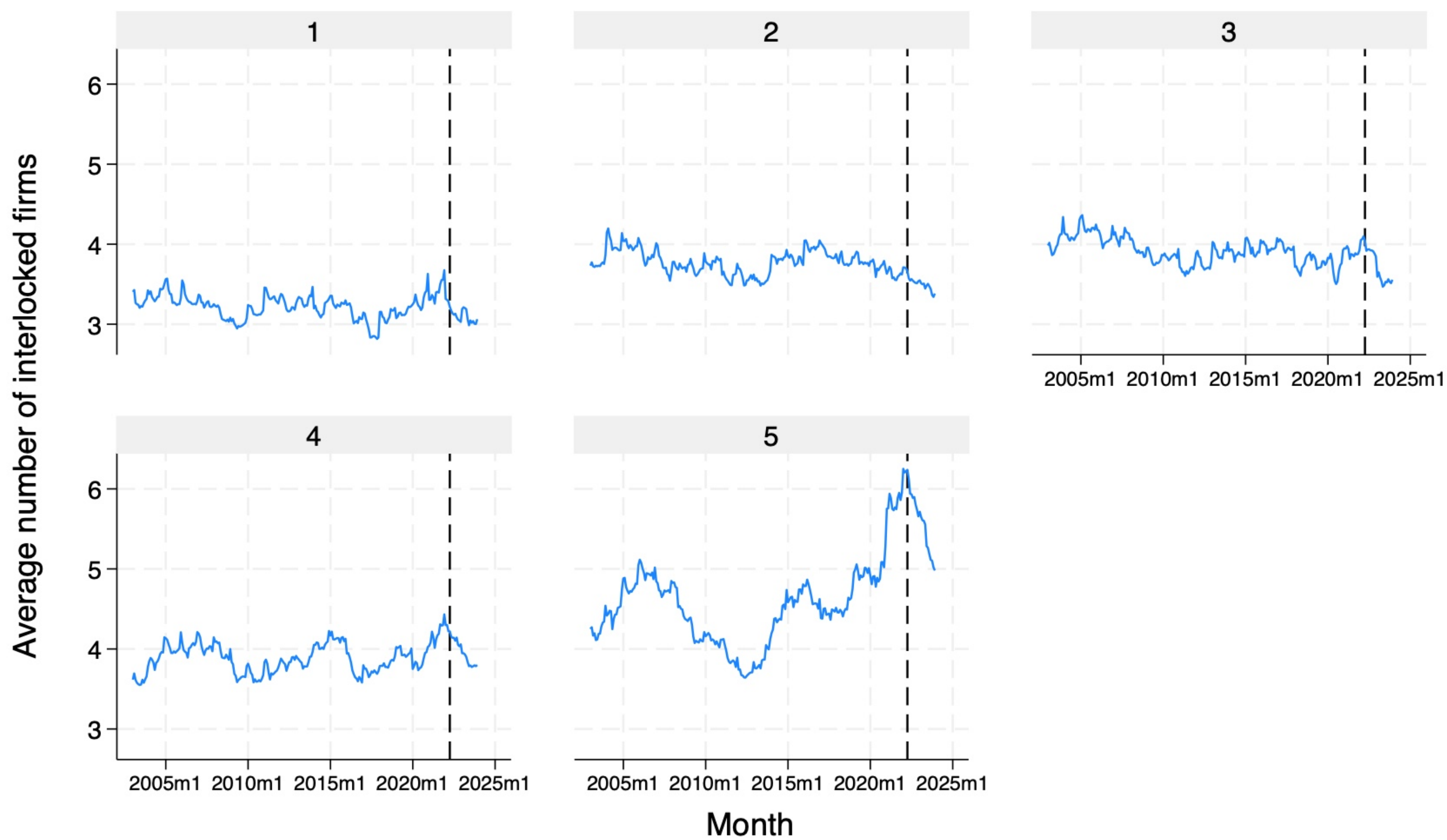
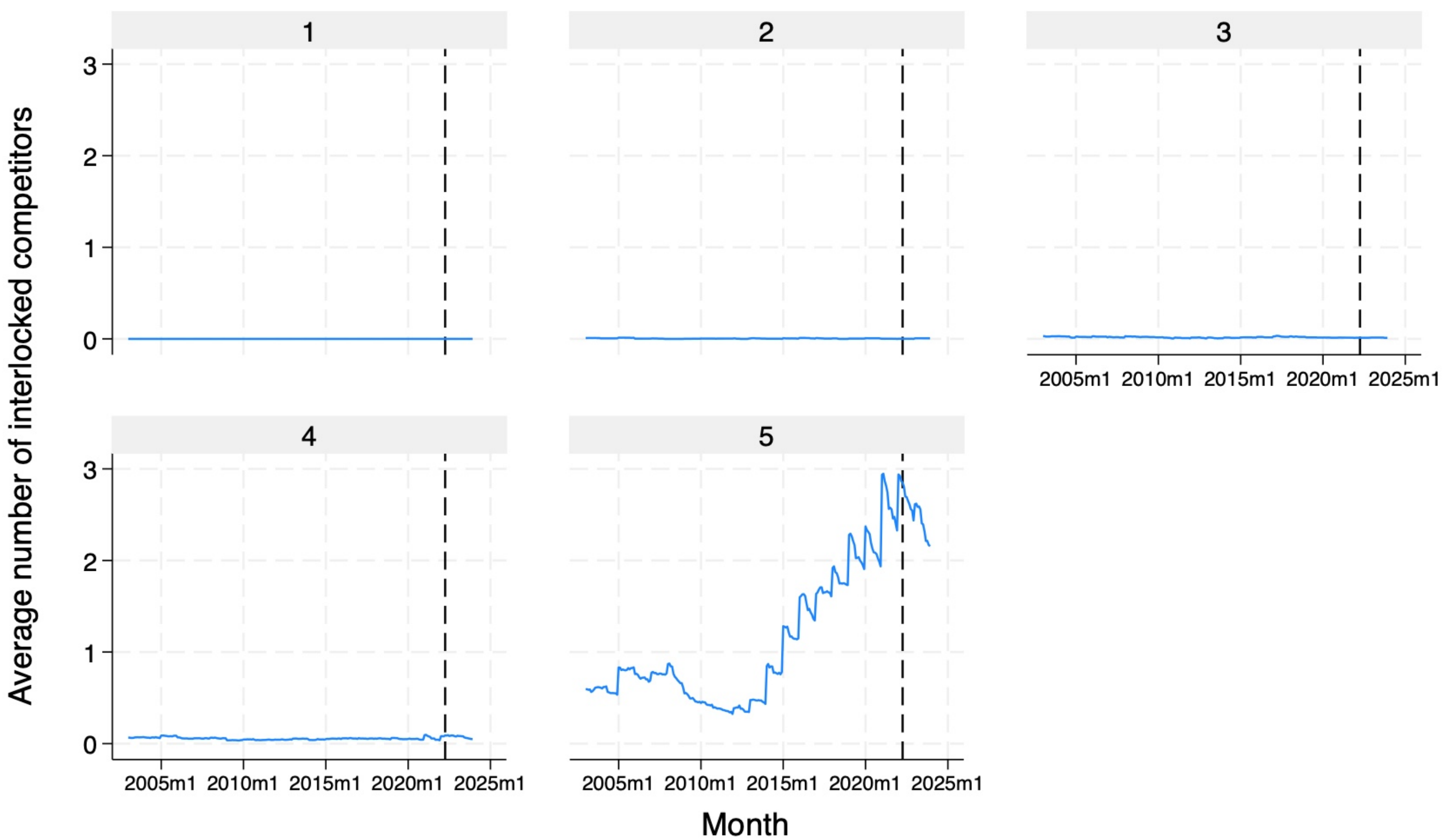


Fig. 6. Evolution of Interlocked Competitors Grouped by TNIC-3 Total Similarity

This plot shows different trends in the number of interlocked product market competitors among groups of stocks with different product similarity with competitors. For each month, I separate all non-financial US public firms into 5 groups based on the firm-month level sum of product similarity scores between the firm and all of its TNIC-3 competitors. Then for each group in each month, I compute the average number of interlocked non-financial competitors across all firms in the group. The y-axis in the plot is the average number of interlocked competitors. The x-axis is the calendar month. The number above each subplot represents the corresponding quintile for each subplot, with 1 denoting the group with the lowest product similarity with competitors and 5 denoting the group with the highest product similarity. Vertical dashed line denotes April 2022, the month of event.



References

Adams, R. B., Ferreira, D., 2007. A theory of friendly boards. The Journal of Finance 62, 217–250.

Adams, R. B., Ferreira, D., 2009. Women in the boardroom and their impact on governance and performance. Journal of Financial Economics 94, 291–309.

Adams, R. B., Hermalin, B. E., Weisbach, M. S., 2010. The role of boards of directors in corporate governance: A conceptual framework and survey. Journal of Economic Literature 48, 58–107.

Amihud, Y., 2002. Illiquidity and stock returns: Cross-section and time-series effects. Journal of Financial Markets 5, 31–56.

Antón, M., Ederer, F., Giné, M., Schmalz, M., 2023. Common ownership, competition, and top management incentives. The Journal of Political Economy 131, 1294–1355.

Barone, G., Schivardi, F., Sette, E., 2025. Interlocking directorates and competition in banking. The Journal of Finance 80, 1963–2016.

Begley, T. A., Haslag, P. H., Weagley, D., 2026. Directing the labor market: Shared directors and employee flows, olin Business School Center for Finance & Accounting Research Paper No. 2024/20.

Burt, A., Hrdlicka, C., Harford, J., 2020. How much do directors influence firm value? The Review of Financial Studies 33, 1818–1847.

Cabezon, F., Hoberg, G., 2026. Leaky director networks and innovation herding. The Review of financial studies 39, 158–197.

Calonico, S., Cattaneo, M. D., Titiunik, R., 2014. Robust nonparametric confidence intervals for regression-discontinuity designs. Econometrica 82, 2295–2326.

Coles, J. L., Daniel, N. D., Naveen, L., 2008. Boards: Does one size fit all? Journal of Financial Economics 87, 329–356.

Dass, N., Kini, O., Nanda, V., Onal, B., Wang, J., 2014. Board expertise: Do directors from related industries help bridge the information gap? The Review of Financial Studies 27, 1533–1592.

Drobetz, W., von Meyerinck, F., Oesch, D., Schmid, M., 2018. Industry expert directors. Journal of Banking & Finance 92, 195–215.

Gao, M., Huang, S., 2024. Independent director tenure and corporate governance: Evidence from insider trading. Journal of financial and quantitative analysis 59, 1760–1795.

Geng, H., Hau, H., Michaely, R., Nguyen, B., 2026. On the purpose of new board overlap enabled by corporate opportunity waivers, working paper.

Geng, H., Hau, H., Michaely, R., Nguyen, B. H., 2025. Common institutional investors and board representation in rival firms. Journal of Corporate Finance 94, 102836.

Gopalan, R., Li, R., Zaldokas, A., 2026. Board connections, firm profitability, and product market actions, european Corporate Governance Institute – Finance Working Paper No. 996/2024.

Harrington, J. E., Skrzypacz, A., 2011. Private monitoring and communication in cartels: Explaining recent collusive practices. The American Economic Review 101, 2425–2449.

Hermalin, B. E., Weisbach, M. S., 1998. Endogenously chosen boards of directors and their monitoring of the CEO. The American Economic Review 88, 96–118.

Herrera-Caicedo, A., Jeffers, J., Prager, E., 2024. Collusion through common leadership, working paper.

Hoberg, G., Phillips, G., 2010. Product market synergies and competition in mergers and acquisitions: A text-based analysis. The Review of Financial Studies 23, 3773–3811.

Hoberg, G., Phillips, G., 2016. Text-based network industries and endogenous product differentiation. The Journal of Political Economy 124, 1423–1465.

Huang, S., Hilary, G., 2018. Zombie board: Board tenure and firm performance. Journal of accounting research 56, 1285–1329.

Intintoli, V. J., Kahle, K. M., Zhao, W., 2018. Director connectedness: Monitoring efficacy and career prospects. Journal of financial and quantitative analysis 53, 65–108.

Lemieux, T., Milligan, K., 2008. Incentive effects of social assistance: A regression discontinuity approach. Journal of Econometrics 142, 807–828.

Masulis, R., Wang, C., Xie, F., Zhang, S., 2025. Directors: Older and wiser, or too old to govern? Journal of financial and quantitative analysis 60, 169–208.

Nguyen, B. D., Nielsen, K. M., 2010. The value of independent directors: Evidence from sudden deaths. Journal of Financial Economics 98, 550–567.

Omer, T. C., Shelley, M. K., Tice, F. M., 2020. Do director networks matter for financial reporting quality? evidence from audit committee connectedness and restatements. Management science 66, 3361–3388.

Poberejsky, R., 2024. Interlocking directorates, competition, and innovation, working paper.

The Rise of Single-Stock ETFs and More Volatile Stock Prices

Abstract

I investigate the novel instrument of single-stock ETFs. I propose that single-stock ETFs gain popularity because they satisfy retail investors' demand for taking short-term leveraged long and short positions on popular stocks at a low cost. I show that flows, turnover ratios, and retail buy-sell imbalances of both long and short single-stock ETFs are positively correlated with retail attention to the underlying stock. This suggests that single-stock ETFs are primarily used by retail investors to circumvent leverage and short-selling constraints. Furthermore, I find that following the launch of the first short single-stock ETF on a stock, the underlying stock experiences a significant and long-lasting increase in idiosyncratic volatility.

1. Introduction

In the last three decades, the ETF industry has been home to many financial innovations, given its ease to trade and its flexible structure. All investors who have access to a brokerage account can trade any ETF listed on the US equity markets whenever the market is open, just as they can trade public stocks. This allows ETF sponsors to directly sell their products to a broader set of potential customers at a lower cost, as compared to traditional mutual fund sponsors. ETFs are also extremely flexible because they are essentially a vehicle that comprises a set of underlying assets, which can be almost any asset specified by the fund sponsor. Despite being initially proposed as equity index-based passive products, ETFs have now been designed to incorporate many strategies, including but not limited to active stock-picking, smart beta, industry-specific, and thematic (Ben-David, Franzoni, Kim, and Moussawi, 2023). However, all these ETFs had one thing in common: they have dozens, if not hundreds, of securities holdings in their portfolios. This changed in July 2022 when a novel type of ETF came to the US market: single-stock ETFs. The portfolios of these single-stock ETFs only have exposure to a single stock. There are two types of single-stock ETFs: long and short. For a long single-stock ETF, it holds a leveraged long position in one underlying stock and promises a daily return that is equal to a pre-specified positive multiple of the daily return of the underlying stock. For a short single-stock ETF, it holds a short position in one underlying stock and promises to deliver a daily return that is a negative multiple of the daily return of the underlying stock. It's important to note that long single-stock ETFs are always leveraged in my data sample, whereas short single-stock ETFs do not necessarily involve leverage. These single-stock ETFs have gained popularity very quickly, as demonstrated by the explosion in their sizes. Within around two years after the introduction of the first single-stock ETF in the US markets, the aggregate total market capitalization of all single-stock ETFs has reached

20 billion dollars in the US. Considering the fact that the aggregate total market capitalization of all single-stock ETFs was 17 million dollars on the first trading day of the first group of single-stock ETFs, the industry grew more than 1000 times in two years. Unlike index-based ETFs that have hundreds or even thousands of portfolio holdings, all single-stock ETFs in my sample data track 24 US stocks in total. In the beginning, the stocks tracked by single-stock ETFs are mainly mega-cap stocks, such as AAPL and TSLA. In 2024 and 2025, the scope of single-stock ETFs expanded to smaller stocks, such as MU, CRWD, and DJT, which are popular among retail investors. If the growth of single-stock ETFs maintains at such a high pace, they will control a sizeable portion of some of the largest stocks in the world, whereas we still do not understand these investment vehicles very well. Regulators are also calling for studies to understand this novel type of instrument[1]. To the author's knowledge, this paper is the first academic research to study single-stock ETFs. With this paper, I wish to fill the gap in our understanding of this financial innovation and its impact on the underlying stocks. Specifically, I aim to answer the following two questions in this paper. First, why single-stock ETFs exist and grow so rapidly. Second, how the launch of single-stock ETFs affects the trading of underlying stock.

Given that single-stock ETFs essentially provide leveraged long and short exposures to individual stocks, the natural question to ask is why investors buy single-stock ETFs instead of buying the stock on margin or short-selling it. To answer this question, I first investigate the underlying mechanisms and some characteristics of single-stock ETFs. I document that single-stock ETFs are expensive. The annual expense ratio for single-stock ETFs averaged 1.13% as of June 2024, which is about three times as high as the average expense ratio for broad-based ETFs, as documented in the literature (Ben-David et al., 2023). I found that most shares (around 83% of market capitalization) of

[1]See SEC commissioner statement at https://www.sec.gov/newsroom/speeches-statements/crenshaw-single-stock-etfs-20220711.

single-stock ETFs are held by individual investors, and the average holding period is as short as two days, which is around 20 times shorter than the average holding period for the index ETFs in my sample period. These results suggest that single-stock ETFs could mainly be held by short-term-oriented individual investors who are willing to pay the hefty fee for some reason. I rationalize their behavior by noting that although single-stock ETFs charge a higher fee than most other ETFs, trading single-stock ETFs is still likely the cheapest, if not the only, way for most individual investors to take on leveraged long or short positions on the underlying stocks.

I further investigate the relationship between single-stock ETF trading and retail investors by focusing on the role of retail investor attention. Adopting the Da, Engelberg, and Gao (2011) retail investor attention measure constructed from the Google Trends search volume index, I show that the flows into both long and short single-stock ETFs are high when the underlying stock receives abnormally high attention from retail investors. The same pattern exists for the turnover ratios of both long and short single-stock ETFs. I also use the Barber, Huang, Jorion, Odean, and Schwarz (2024) algorithm to directly identify retail buy and sell trades of the single-stock ETFs and compute a retail investors' buy-sell imbalance measure, similar to that in Barber and Odean (2008). I show that when retail attention on the underlying stock increases, the buy-sell imbalance ratio of retail investors increases for both long and short single-stock ETFs. Moreover, for long single-stock ETFs, both attention-driven buying and attention-driven selling from retail investors are positive and significant. These findings fit the limited attention theory of retail investors (Seasholes and Wu, 2007; Barber and Odean, 2008; Barber, Huang, Odean, and Schwarz, 2022): Retail investors do not always attend to information about public firms and financial markets, and they are more likely to make trades when their attention is drawn to the stocks. The evidence discussed so far points to an answer to the question of why single-stock ETFs exist and grow: these ETFs cater to retail investors' demand for short-term leverage and short-selling by providing them with a cheap and accessible way to do so.

After answering the first question, I move on to investigate how the launch of single-stock ETFs affects the trading of the underlying stocks. In this paper, I focus on the idiosyncratic volatility of the underlying stocks because if retail trading is driven by attention rather than risks, more retail trading should lead to higher volatility in stock returns after controlling for risk factors. This hypothesis is similar to the story of Ben-David, Franzoni, and Moussawi (2018), where they show that ETFs attract a new group of short-term traders, resulting in an increase in the volatility of the stocks in the ETF basket. I compute the idiosyncratic volatility for each stock in each month as the standard deviation of the residuals from the CAPM regression. I find that following the launch of the first short single-stock ETF on an underlying stock, the idiosyncratic volatility of the stock increases, and this increase lasts for at least 12 months. Following the launch of the first long single-stock ETF, the increase in the stock idiosyncratic volatility is positive but insignificant, and the increase reverts to zero in the eighth month after the long ETF launch.

This article is mainly related to three topics in literature: ETF innovation, impact of ETFs on underlying assets, and retail trading. First, this article contributes to the ETF innovation literature by investigating a new type of ETF and explaining its rapid growth since initial launch. Many authors have attempted to explain the rise of non-index-based ETFs. Cong, Huang, and Xu (2024) argue that ETFs that are not fully index-based allow investors to use their factor-specific information to make trading profits. They suggest that in equilibrium, the weight of each asset in the portfolio should be proportional to its exposure to a factor to facilitate "factor investing". Ben-David et al. (2023) investigate the rise of sector-based and thematic ETFs and suggest that these ETFs mainly target sectors that have high investor attention and sentiment in recent months. My argument is in a similar spirit to that of Ben-David et al. (2023). I show that another type of novel ETF, single-stock ETF, is also a popular vehicle that facilitates trading of attention-grabbing stocks. I argue that single-stock ETF relaxes the leverage and short-selling constraints faced by retail investors when trading attention-grabbing stocks, resulting in single-stock ETFs' popularity and rapid growth. While Bessembinder (2025) also investigates the leveraged single-stock ETFs, Bessembinder (2025) focuses on the underlying mechanisms for multi-day returns of these instruments, whereas I mainly investigate the clienteles of these instruments and their impact on underlying stocks. Second, this article contributes to the literature on the impact of ETFs on underlying assets. Scholars have shown that ETFs make underlying stock returns more correlated (Da and Shive, 2018), and widen the bid-ask spread of underlying stocks (Evans, Moussawi, Pagano, and Sedunov, 2026). There has also been conflicting evidence on how ETFs affect the informativeness of underlying stock prices (Israeli, Lee, and Sridharan, 2017; Buss and Sundaresan, 2023). There have been articles showing that ETFs raise the volatility of underlying stocks (Ben-David et al., 2018; Cheng and Madhavan, 2009). In this paper, I add support to this strand of literature by demonstrating that following the introduction of short single-stock ETFs, there is a long-lasting increase in the idiosyncratic volatility of the underlying stock. This finding could be explained by the theory that single-stock ETFs attract attention-driven retail investors, whose trading adds noise to the underlying stock price. Third, this article contributes to the literature of retail investor trading by exploring a new instrument that could serve as a useful arena for observing retail behaviors. I show that trading in single-stock ETFs exhibits many patterns that are typical of retail trading, suggesting that the majority of investors trading these single-stock ETFs are retail investors. In particular, I apply the retail attention measure from Da et al. (2011) to single-stock ETFs and show that this measure predicts trading in single-stock

ETFs. I also confirm the finding of Barber and Odean (2008) that attention-driven retail buying is greater than attention-driven retail selling, in the context of single-stock ETFs. Moreover, I find that both long and short single-stock ETFs experience similar magnitudes of increase in retail buy-sell imbalance when retail attention is high. This supports the explanation proposed in Barber and Odean (2008) that asymmetry in attention-driven buying and selling is due to the short-selling constraint of retail investors, instead of because retail investors are more likely to long the stock when attention is high than to short the stock.

The remainder of the paper is organized in the following way. Section 2 introduces the sources of the data I use in this paper. Section 3 provides a general overview of single-stock ETFs. In section 4, I investigate the relationship between retail investor attention and single-stock ETF trading. Section 5 examines how the idiosyncratic volatility of the underlying stock evolves around the launch of the first single-stock ETFs. Section 6 concludes the paper.

2. Data

First of all, there is no complete list of all single-stock ETFs provided by any data vendor. Therefore, I manually read through the list of all US dollar-denominated ETFs that have ever been traded (totaling 9344 ETFs as of March 2025) from Refinitiv Eikon. Because this list also includes the ETFs that have stopped trading, there is no survivor bias in my final sample of single-stock ETFs. I identify all single-stock ETFs by the inclusion of any stock ticker in the name of the ETF. For example, "GraniteShares 2x Short NVDA Daily ETF" is a short single-stock ETF on the stock NVDA. It is a convention that single-stock ETFs need to include the ticker of the underlying stock in their names. However, there are many ETFs with stock tickers in their names that are not the standard single-stock ETFs I want to focus on in this paper. Specifically, there are many option income strategy ETFs, which hold publicly traded stock options to create a hedged portfolio that generates streams of monthly income for investors. These option income ETFs mostly adopt covered option strategies and thus have a very different underlying mechanism from the single-stock ETFs that I wish to study in this paper; therefore, I exclude these option income ETFs from my sample. I further exclude the ETFs that have never been traded on any US exchanges by the end of 2024. After this screening process, I am left with 72 single-stock ETFs on 24 underlying stocks. Out of the 72 single-stock ETFs, 47 are long ETFs, and 25 are short ETFs. Note that only 66 out of 72 single-stock ETFs were still being traded as of December 2024. The list of the names and tickers for all single-stock ETFs in my sample can be found in Appendix C. The earliest trading day of any single-stock ETF was July 13, 2022. Therefore, the sample period for single-stock ETFs is from July 13, 2022, to December 31, 2024.

I download all daily data on single-stock ETFs from CRSP. These daily data include price, return, trading volume, net asset value, etc. However, the daily net asset value data from the CRSP mutual fund dataset contain many missing values for some single-stock ETFs; thus, I focus on the market capitalization of these ETFs as the main measure of size. This measure of size is warranted by the small daily price premia in these single-stock ETFs, computed from the limited net asset value data. Moreover, I download the daily data on underlying stocks from CRSP. Quarterly institutional ownership data on single-stock ETFs are from the Thomson Reuters 13-F database. Earnings announcement dates for stocks are accessed from IBES. I use Bloomberg to download the cross-sectional data on fees and expense ratios for single-stock ETFs as of June 2024. I am unable to download panel data on ETF fees and expense ratios from CRSP due to the incompleteness of such data on CRSP. As a result, I am unable to include ETF fees and expense ratios as control variables in my panel regressions presented later in the paper. The fees and expense ratios for single-stock ETFs will be used for informational purposes only. Following the literature (Boehmer, Jones, Zhang, and Zhang, 2021; Barber et al., 2024), I download trade-level data from NYSE Trade and Quote (TAQ) to measure retail trading of single-stock ETFs.

Lastly, I download historical Google Search Volume Index (SVI) from Google Trends, as in Da et al. (2011). Following Da et al. (2011), I manually search the stock ticker for each of the underlying stocks. For example, for "Apple Inc.", I query for the keyword "AAPL" on Google Trends. As argued by Da et al. (2011), the practice of searching for stock tickers is less ambiguous because people searching for the stock ticker are very likely to be attentive to the information about the stock, which is exactly the attention I want to measure. Da et al. (2011) find strong empirical support for the reliability of this retail attention measure. I construct two time series of retail attention measure (Abnormal SVI, or ASVI) for each underlying stock at two different frequencies. First, I construct the weekly time series of SVI in the same way as Da et al. (2011). The query time period is from Jan 1st, 2020 to Dec 31st, 2024. I limit the geographical region to the United States. Then I query for the ticker symbols for the underlying stocks to get the weekly SVI data[2]. Second, I construct the daily time series of SVI. Because any query with a period longer than a few months will return weekly SVI values, rather than daily SVI, I write a scraper program to query the stock ticker for 31 days at a time. There are two caveats with Google Trends SVI data that will affect my variable preparation. The first caveat is that for each query, Google automatically performs linear rescaling on the SVI data such that the maximum value in the returned time series is always 100. The second caveat is that SVI values are rounded to the nearest integer. To mitigate the misaligned scales and rounding errors resulting from

[2]Google Trends performs random sampling to produce SVI index for each query, and thus the results will be slightly different for the same query parameters if the query is repeated. Da et al. (2011); Cebrián and Domenech (2024) show that such sampling errors are small.

these two caveats, I leave a 5-day overlap window between each pair of adjacent query periods for the same stock. Then I re-align the scales of the SVI series returned by two adjacent queries by linearly scaling the two series such that the greatest SVI value in the 5-day overlap window for each series is equal. I choose to align the scale based on the greatest value in each series in the overlap window to minimize rounding errors. For example, suppose I only have a 1-day overlap and align the SVI series by setting the SVI value on that overlapping day equal in both series. If the actual precise SVI value on that overlap day is 1.4 in the earlier series and 1.5 in the later series, then the SVI value after rounding will be 1 in the earlier series and 2 in the later series. Because I can only access the rounded SVI value, I would have to scale all observations in the earlier series by a factor of 2, whereas the actual scaling factor should have been 1.07 (1.5/1.4). As one can see, the rounding error will be very significant in this case. To minimize the rounding error, I choose to adopt a 5-day overlap period and scale the series based on the greatest value in that window. While one can argue for other lengths of the overlapping window, I choose 5 days to strike a balance between the rounding error minimization and the total scraping time minimization. Similarly, using a query period of 31 days at a time is also the result of balancing these two problems. Using a longer query period will reduce the time needed for scraping, whereas it will compress the SVI values in the series downwards, because the maximum value is always 100. After setting the daily SVI values to the same scale for each stock, I concatenate them together and perform the final rescaling such that the maximum daily SVI value is 100 for each underlying stock. Now I have got a daily panel of SVI measure for each underlying stock from Jan 1^{st}, 2020 to Dec 31^{st}, 2024, and each time series is on the same scale. Then, I only keep daily SVI observations on the trading days. Note that, for some tickers that went public in the middle of my sample period, Google Trends does not give any SVI values before the firm went public. This makes sense, because the character combinations in many stock tickers were meaningless before the stock went public.

The weekly SVI data from Google Trends correspond to each week from Monday to Sunday. There is concern that the retail attention reflected in Google searches over the weekend will not impact the trades in the same week. This bias will significantly affect the construct validity of the weekly SVI index only if the SVI index is high on the weekend. However, an inspection of the daily SVI index will mitigate this concern because the daily SVI measure on Saturdays and Sundays is usually close to zero. Therefore, the weekly SVI measure I obtain mainly captures retail investor attention during the workdays of the week, and most workdays are trading days when investors can trade stocks and ETFs. After obtaining SVI measures on both daily and weekly frequencies, I need to compute abnormal SVI (ASVI) values as proxies for retail attention. As the first step, I subtract the median of weekly SVI values from the past eight weeks from the current weekly SVI, as in Da et al. (2011). For daily SVI. I subtract the median daily SVI from the past 42 trading days from the current daily SVI. Following Da et al. (2011), I then standardize each series at both weekly and daily frequencies obtained from the previous step to obtain the final retail attention measure (Abnormal SVI, or ASVI).

Because I download SVI measures ticker-by-ticker, there is no cross-sectional comparability between SVI values for different underlying stocks. This is the rationale behind standardization by each stock. The daily ASVI values for each underlying stock are plotted in figure 1. There are regular spikes in the ASVI for most stocks, and I confirm that almost all spikes occur around the time of quarterly earnings announcements. There is a concern that ASVI might also simultaneously capture some institutional attention, because institutions are also attentive to company earnings announcements. I will address this concern in the next section, and I will show that the main contributor to stock ASVI and single-stock ETF trading is retail investors. The spikes in ASVI could also raise concern that the regression results in the following sections might be driven by outliers around earnings announcements. To alleviate this issue, I conduct (unreported) robustness tests by running all the following regressions again, but with 10 daily observations around each earnings announcement date dropped from the sample. The results from these robustness tests are qualitatively similar to the regression results reported in the following sections.

3. Overview of Single-Stock ETFs

This section introduces the underlying mechanisms, institutional backgrounds, and characteristics of single-stock ETFs in greater detail.

3.1. Underlying Mechanisms of Single-Stock ETFs and Implications

All single-stock ETFs in my sample aim to provide a pre-specified multiple of the daily returns of the underlying stock by using derivatives. This pre-specified multiplier can be positive or negative. If the multiplier is positive, it is called a long ETF. If the multiplier is negative, it is a short ETF. The name of a single-stock ETF always contains information about both its multiplier and its direction. Long ETFs can be identified by relevant keywords, such as "Long" and "Bull", in the ETF names. The names of short ETFs usually contain keywords like "Short", "Bear", and "Inverse". For example, "GraniteShares 2x Short NVDA Daily ETF" is a short ETF tracking Nvidia stock with a multiplier of -2. If the multiplier is positive, it is always greater than 1. In other words, long ETFs are always leveraged. There is no value for investors in holding a long single-stock ETF with a multiplier of 1, because holding such an ETF is always inferior to holding the underlying stock. On the other hand, the leverage ratios of all short single-stock ETFs in my sample range from 1 to 2. This implies that some investors are willing to buy unlevered short single-stock ETFs, possibly because they face short-sale constraints or high short-selling fees.

Single-stock ETFs track a multiple of the daily return of underlying stocks by signing swap agreements with a group of large investment banks. At the end of each trading day, the management team of a single-stock ETF adjusts the notional value of the swaps to maintain the correct level of multiplier for the next trading day. This practice is known as "daily rebalancing". All single-stock ETFs in my sample perform daily rebalancing; thus, they are required by the SEC to include the word "daily" in their names[3]. This daily rebalancing feature might also be a reason for the popularity of these single-stock ETFs, because it would be inconvenient, if not impossible, for retail investors to accurately rebalance their portfolios every day. The costs and benefits of daily rebalancing to investors will be an issue worth future exploration. Conventional index-based leveraged and inverse ETFs also perform daily rebalancing through financial derivatives. Over an extended investment horizon, daily rebalancing makes the holding period return of leveraged and inverse ETFs significantly deviate from the specified multiple of the holding period return of the underlying stock or index. This brings a nuanced difference between holding a leveraged ETF and margin trading. Suppose two investors achieve a leverage of x ($x > 1$) on the same stock through buying-on-margin and long single-stock ETFs respectively. Each of them has an endowment M. Both investors hold the position for T trading days. The holding-period return for the investor who buys the stock on margin, assuming zero interest, will be given by:

$$r_{Margin,T} = \frac{xM\prod_{t=1}^{T}(1+r_t) - xM}{M} = x(\prod_{t=1}^{T}(1+r_t) - 1) \tag{1}$$

[3]See question 20 at SEC correspondence https://www.sec.gov/Archives/edgar/data/1587982/000139834422010342/filename1.htm

where r_t is the underlying stock return on day t.

The holding-period return for the investor who buys a long single-stock ETF will be as follows:

$$r_{ETF,T} = \frac{M \prod_{t=1}^{T}(1 + xr_t) - M}{M} = \prod_{t=1}^{T}(1 + xr_t) - 1 \tag{2}$$

The holding-period return for holding the underlying stock without leverage is:

$$r_{Stock,T} = \prod_{t=1}^{T}(1 + r_t) - 1 \tag{3}$$

We can see that for a holding period $T > 1$, the return from buying-on-margin will be different from the return of holding a single-stock ETF. This difference will be larger for longer holding periods. In summary, buying the stock on margin yields an accurate multiple of the holding period stock return, whereas buying single-stock ETFs provides an accurate multiple of the daily stock return to the investor every day.

3.2. Institutional background

Single-stock ETFs were introduced in Europe slightly earlier than in the US. The first single-stock ETF was introduced to the US market in 2022. The launch in 2022 may be partly due to a 2021 SEC regulation[4] that modified rules 6c-11 and 18f-4 in the Investment Company Act and significantly reduced the regulatory burden on leveraged and inverse ETFs. Under the old regulation, every leveraged and inverse ETF needed to obtain an exemptive order from the SEC before it could be introduced to the market. Under the new regulation, leveraged and inverse ETFs are automatically exempted, significantly lowering the costs and uncertainty associated with launching new leveraged and inverse ETFs. ETF sponsors can now launch leveraged and inverse ETFs much

[4] Accessed at https://www.sec.gov/files/rules/final/2019/33-10695.pdf

quicker and cheaper than before. Since the introduction of the first single-stock ETF in the US, the aggregate total market capitalization for all single-stock ETFs has been expanding rapidly, as illustrated in figure 2. Within around two years, the aggregate total market capitalization increased from 0 to more than 20 billion dollars.

3.3. Characteristics of Single-Stock ETFs and Underlying Stocks

Table 1 summarizes the key characteristics of single-stock ETFs. The first four rows are summary stats at ETF-daily level. As we can see from table 1, the ETF-daily level average total market capitalization is around 143 million dollars. Note that the total market capitalization of almost every single-stock ETF increases throughout the sample period, and thus the average total market capitalization at the end of the sample period is much higher than the sample mean in table 1. On average, long ETFs are roughly 7.5 times larger in size than short ETFs. The size discrepancy between long and short single-stock ETFs grew throughout the sample period. One of the most important results in the summary statistics is that the daily turnover ratios are very high for both long and short ETFs. The investors' average holding period for long ETFs is about 4 days. For short ETFs, investors only hold them for around 1.2 days on average. As a comparison, the average daily turnover ratio for all ETFs in the CRSP dataset in the same sample period is 2.08%. The 90-th percentile of average ETF daily turnover ratio is 3.20%. Average daily turnover ratios for single-stock ETFs are more than ten times higher than the average for all ETFs. This observation points to the insight that most investors of single-stock ETFs use them as a short-term trading tool, rather than a long-term investment vehicle. I compute daily proportional flows of single-stock ETFs in a similar way as in Sirri and Tufano (1998). On average, long single-stock ETFs acquire 2% inflow every day in my sample, while short single-stock ETFs enjoy a slightly higher inflow of around 3% every day. These daily average proportional flows are large, and corroborate with the rapidly growing sizes of single-stock ETFs. The standard deviations of these proportional flows are also very high, with 17% for long ETFs and 132% for short ETFs, signifying that daily flows are very volatile. These patterns hold true for daily dollar flows of single-stock ETFs. Single-stock ETFs are usually traded at a price that's very close to their net asset values, as illustrated by the small premium. This suggests that market makers of these single-stock ETFs excelled in arbitraging and providing liquidity to these instruments.

At the quarterly level, we can see that most of the shares (83%) of single-stock ETFs are held by individual investors. Figure 3 further shows that the institutional ownership percentage is decreasing over time for both long and short single-stock ETFs. This pattern remains true for most of the individual single-stock ETFs. However, because the total market capitalization of most of the single-stock ETFs is increasing rapidly, institutions are increasing the value of their single-stock ETF holdings, despite the drop in the ownership percentage. These ownership data are recorded at the end of each calendar quarter. There is no daily data on institutional ownership, and the quarterly ownership data may not reflect the actual institutional ownership within the quarter.

The last six rows in table 1 are at ETF-level. For each single-stock ETF, I compute its tracking error using two measures. Elton, Gruber, and Busse (2004) propose two types of tracking error for index funds in their case. The first type of tracking error is the random deviation of fund performance from underlying index performance due to imperfect replication. They measure this type of tracking error using 1 minus R-squared for the following daily level time series regression:

$$r_{t,NAV} = \alpha + \beta \times r_{t,index} + \epsilon_t \tag{4}$$

where $r_{t,NAV}$ is the daily percentage change in the fund's net asset value on day t, and $r_{t,index}$ is the daily percentage return of the target index of the fund. I replicate this measure by regressing the daily percentage change in net asset value of each single-stock ETF on the daily percentage return of the underlying stock. In my sample, the average R-squared is around 99% for both long and short single-stock ETFs. This R-squared is slightly smaller than the average R-squared (99.99%) for all S&P 500 index mutual funds in Elton et al. (2004). Therefore, single-stock ETFs have larger random variations in their net asset values than S&P 500 index funds. This tracking error may be attributed to the use of derivatives in managing these ETFs, as opposed to holding underlying stock baskets in most index mutual funds. The second type of tracking error proposed by Elton et al. (2004) is the systematic deviation of net asset value from the underlying return. They capture this type of tracking error by using $|1 - \hat{\beta}|$, where $\hat{\beta}$ is the point estimate of the coefficient in the above regression equation. This measure is not applicable to single-stock ETFs, because all single-stock ETFs are levered, inverse, or both. Therefore, desired β for all single-stock ETFs will not be 1. Hence, I modify my second tracking error measure to capture the spirit of systematic deviation of net asset value, as envisioned by Elton et al. (2004). I run the following daily regression for each single-stock ETF:

$$r_{t,NAV} - k \times r_{t,stock} = \alpha + \beta \times r_{t,stock} + \epsilon_t \tag{5}$$

where k is the pre-specified return multiplier for this single-stock ETF. This way, I can use the estimate for the coefficient β to measure the systematic deviation of the ETF net asset value return from the pre-specified multiple of the underlying stock return. Running regressions in this form deals with the issue above such that my measure is applicable to all single-stock ETFs, despite their wide range of multipliers. Unlike in Elton et al. (2004), the $\hat{\beta}$ in the above regression should have been 0, if there is no systematic tracking error. Therefore, I use $|\hat{\beta}|$ as the measure for systematic tracking error. From table 1, one can see that long single-stock ETFs have slightly higher systematic tracking errors than short single-stock ETFs, both of which have higher systematic tracking errors than S&P 500 index funds (Elton et al., 2004).

Expense ratios and management fees data were downloaded from Bloomberg terminal in early June 2024. Due to data limitation of Bloomberg terminals, these expense and fee data are all cross-sectional as of the date of download. Thus, expense ratios and management fees are not included as control variables in panel regressions in the following sections. All long and short ETFs have similar expense ratios and management fees. On average, single-stock ETFs charge 1.13% of their total net asset every year. The average management fee of single-stock ETFs is 0.93%. These fees are extremely high compared to the fees charged by index ETFs, which commonly have annual expense ratios at around a few basis points. One of the possible reasons for their high expense ratios is that single-stock ETFs perform daily rebalancing, which demands the use of expensive swaps. The important thing to keep in mind is that despite the high costs, trading single-stock ETFs is still likely the cheapest and the most convenient way for retail investors to leverage or short-sell many stocks, compared to other means[5]. Finally, both long and short single-stock ETFs have average leverage ratios greater than 1, and long single-stock ETFs have higher leverage ratios than short single-stock ETFs on average. These findings are consistent with the earlier observation that all long single-stock ETFs are levered, while some short single-stock ETFs do not use leverage.

In addition to characteristics of single-stock ETFs, I also present the characteristics of all underlying stocks of single-stock ETFs in table 2. There are 24 US stocks that ever had a tracking single-stock ETF as of the end of December 2024. On a daily level, the average market capitalization of these 24 underlying stocks is 480 billion dollars, indicating that most of the underlying stocks are mega-cap stocks. The average daily

[5] See Robinhood margin trading and short-selling policy (https://robinhood.com/us/en/support/articles/margin-rates/)

stock return is 0.15% in my sample. The stock turnover ratio is 1.98% on average, much smaller than the turnover of single-stock ETFs. The average book-to-market ratio for underlying stocks is 0.27, implying that most of the underlying stocks fall into the growth category. I compute the daily Amihud illiquidity measure for stocks by scaling the daily absolute return in decimal by daily trading volume in billion dollars. On an average day, the underlying stock price moves by 9% for every billion dollars traded in the stock. The last four rows in table 2 compare the size of single-stock ETFs and the size of underlying stocks. I aggregate all the long or short single-stock ETFs tracking a specific stock to stock-daily level, and then compare it to underlying stocks. The aggregate market capitalization of both long and short single-stock ETFs is still very small compared to the market capitalization of underlying stocks, as most of the underlying stocks are large stocks. Despite the relatively small market capitalization of single-stock ETFs, the dollar trading volumes of single-stock ETFs are significant relative to the trading volumes of underlying stocks. On average, the dollar amount of long single-stock ETFs traded every day stands around 0.84% of underlying stock trading volume, while the trading volume of short single-stock ETFs is 0.17% of stock trading volume. More importantly, this ratio is highly positively skewed, especially for long ETFs. The 90th percentile of this ratio for long single-stock ETFs is 1.97%, and there are dozens of daily observations of this ratio above 10%, with the maximum standing around 20%. This suggests that on some days, the amount of capital trading single-stock ETFs is very significant compared to the underlying stocks. Therefore, it's very likely that the arbitraging demand between single-stock ETFs and underlying stocks may have a huge impact on the underlying stock price, and the leveraged nature of single-stock ETFs only adds to this impact even further.

In sum, results in tables 1 and 2 allude to a story that single-stock ETFs are mainly used by short-term oriented individual investors to take on leveraged or inverse positions on some of the largest and most popular stocks in the US market. This story could explain the emergence and the rapid growth of single-stock ETFs in the last two years. To provide further support for this story, I attempt in the next section to show that there is significant retail trading in single-stock ETFs.

4. Single-Stock ETF Trading and Retail Investor Attention

The results in the last section point to the idea that single-stock ETFs exist to facilitate short-term leveraging and short-selling by retail investors. If this story is true, then single-stock ETFs should be traded mainly by retail investors. In this section, I am going to test this hypothesis by focusing on retail attention. I test how retail attention on the underlying stock is related to three characteristics of single-stock ETFs: flows, turnover ratios, and retail trades.

4.1. Single-Stock ETF Flows and Retail Attention

I resort to the investor attention literature (Seasholes and Wu, 2007; Barber and Odean, 2008; Barber et al., 2022) to uncover the relation between retail attention and single-stock ETF trading. Literature has shown that retail investors have the tendency to trade attention-grabbing stocks. Therefore, I will test whether retail investor attention on the underlying stocks positively affects the turnover ratios, flows, and retail trades of single-stock ETFs in this section.

I follow Da et al. (2011) to measure retail investor attention using Abnormal Search Volume Index (ASVI) from Google Trends. They have shown that this measure captures the attention of retail investors and correlates with their trading behaviors. Da et al. (2011) also argue that this ASVI measure is superior to other common investor attention measures, such as extreme stock turnover and returns, because it is more "active" than other measures. Even if a stock has extreme turnover or return on one day, it's not guaranteed that this extreme phenomenon will be noticed by many retail investors, since retail investors usually do not pay attention to the stock market every day. In contrast, searching on Google can only happen after investors are already attentive to the stock. Therefore, Google search volume is a less noisy measure for the attention of retail investors than common trading-based measures.

First, I investigate the relation between retail investor attention and single-stock ETF flows. I measure the daily (weekly) proportional flow of each single-stock ETF by using the formula from Sirri and Tufano (1998), except that I use the market capitalization of ETFs, instead of total net assets:

$$Flow_t = \frac{MktCap_t - MktCap_{t-1} \times (1 + r_t)}{MktCap_{t-1}} \tag{6}$$

where $MktCap_t$ is the market capitalization of the single-stock ETF at the end of period t, and r_t is the price return of this ETF in period t. I use the market capitalization of ETFs, instead of the total net assets of ETFs, mainly because CRSP reports many missing daily observations of total net assets for single-stock ETFs, while very few observations of market capitalization are missing.

I run the following weekly panel regression to investigate the relation between retail attention on the underlying stock and proportional flow of the single-stock ETF:

$$Flow_{it} = \alpha + \beta_1 \times ASVI_{i,t-1} + \beta_2 \times Controls_{i,t-1} + \delta_t + \gamma_i + \epsilon_{it} \tag{7}$$

where $ASVI_{i,t-1}$ is the standardized abnormal search volume index of the underlying stock of single-stock ETF i in week $t-1$. I use lagged stock ASVI as my main independent variable, because using concurrent retail attention would make it difficult to interpret the direction of the effect, even if β_1 is significant. Control variables include other common retail attention measures and the variables known to affect ETF flows. Researchers (Barber and Odean, 2008; Barber et al., 2022) have shown that stocks with high absolute value of return and high turnover ratio will attract more retail investor attention, which will induce them to trade the stocks more. Highly volatile daily price movements usually make the news and attract investor attention. I use ASVI as my main variable of interest, because Da et al. (2011) show that ASVI is a better investor attention proxy than these control variables, in that ASVI contains more useful information about investor attention. Moreover, Da et al. (2011) provide evidence that ASVI measures the attention of retail investors specifically, who are the main group of investors related to my story. It's well-known that fund return, fund size, and fund age are related to the fund flow (Sirri and Tufano, 1998; Spiegel and Zhang, 2013), so I include these as controls. All control variables are lagged by 1 week, because most of these variables are related to the trading of underlying stocks, and thus can potentially be affected by the trading of single-stock ETFs in the same week through the arbitrage channel. The standard error is double-clustered at the ETF and week level. I run the above regression on all ETFs, long ETFs, and short ETFs respectively. The output of the regression above is presented in table 3.

The main observation in table 3 is that the weekly flows of both long and short ETFs are positively predicted by the retail attention in the previous week. This means that both long and short single-stock ETFs experience higher weekly flow when the underlying stock attracts higher attention from retail investors. Another finding in table 3 is that sizes of long and short ETFs are negatively related to the ETF flows, even after controlling for retail attention. This observation is consistent with the traditional findings in mutual fund literature (Sirri and Tufano, 1998).

My hypothesis is that retail attention on the stock induces flow into the corresponding single-stock ETFs, which is supported by findings in table 3. Nevertheless, one might expect that retail attention is short-lived and does not last a week. This suggests that the effect I found in table 1 should also be present at higher frequencies. Thus, I rerun the regression on daily frequency with lagged ASVI:

$$Flow_{it} = \alpha + \beta_1 \times ASVI_{i,t-1} + \beta_2 \times Controls_{i,t-1} + \delta_t + \gamma_i + \epsilon_{it} \tag{8}$$

where all variables are at daily level, and all independent variables are lagged by 1 day.

The output of this regression is presented in table 4. Because the regression is at daily level, I do not have stock volatility as a control variable. As seen in table 4, ASVI of the underlying stock on the previous day positively predicts the flow into both long and short single-stock ETFs, despite the positive relation being insignificant for short ETFs this time. This daily-level result from table 4 corroborates with the weekly-level result from table 3, in that both tables suggest the flows of both long and short single-stock ETFs are positively related to retail attention on the underlying stock of the ETFs.

4.2. Single-Stock ETF Turnover and Retail Attention

Besides ETF flow, turnover ratio is also an important indicator of the intensity of ETF trading. I further investigate the relation between stock retail attention and ETF turnover ratio. I run the same regression as above, with the dependent variable changed to the turnover ratio of the single-stock ETFs. Again, I run this regression on both weekly and daily levels. The output of the weekly regression is shown in table 5, and the daily results are in table 6. The results for turnover ratios are qualitatively identical to the results for fund flows. On the weekly level, the lagged retail attention on the underlying stock positively predicts the turnover ratios of both long and short single-stock ETFs in the following week. In daily panel regressions, the lagged retail attention on the underlying stock positively predicts the turnover ratios of both long and short single-stock ETFs on the following day.

Tables 3, 4, 5, and 6 demonstrate that flows and turnover ratios of both long and short single-stock ETFs are positively related to the retail investor attention on the underlying stock. This finding corroborates my story that single-stock ETFs are vehicles mainly traded by retail investors. Note that the flows and turnover ratios of both long and short single-stock ETFs are symmetrically influenced by investor attention. This symmetric influence on trading of ETFs in both directions is different from the finding of Barber and Odean (2008). Barber and Odean (2008) show that retail investors are net buyers of attention-grabbing stocks, and they explain this phenomenon with the short-sale constraint faced by retail investors. They argue and empirically show that retail investors only sell the stocks that they already own. In the case of single-stock ETFs, since a short equity position on the underlying stock is achieved by buying short ETFs, retail investors' short-selling behavior is no longer restricted by the short-sale constraint posed by their brokers. Retail investors can buy both long and short single-stock ETFs when their attention is drawn to the underlying stock, as shown in tables 3 and 4. As a result, despite differing from the empirical findings of Barber and Odean (2008) on face value, my findings above are consistent with and provide support for their investor attention theory. Furthermore, my results allude to that short single-stock ETFs effectively relax short-selling and leverage constraints on retail investors, allowing them to magnify their influence on the underlying stocks, especially when the underlying stocks are smaller in size.

I have interpreted the positive relations between retail attention and ETF flow and turnover as supporting evidence for the retail trading story. However, one may challenge my story by asserting that Google search volume is correlated to fundamental news and thus is also related to institutional trading. I would like to refute this alternative hypothesis both logically and empirically. Logically, I want to make two points to argue that this alternative hypothesis is not likely to be true. First, Da et al. (2011) establish that Google search volume mainly captures the attention of uninformed retail investors. Their reasoning is that unlike retail investors who tend to use Google, institutional investors use more sophisticated tools, such as Bloomberg terminals, to gather information. They further segregate retail investors into informed and uninformed groups. They empirically show that the Google search volume index correlates more to the order flows of uninformed retail investors than to the orders of informed retail investors. Therefore, the retail attention measure I use in my analysis is not likely to be correlated to informed retail trading, not to mention institutional trading. Second, the institutional trading story cannot explain the empirical finding that flows of both long and short single-stock ETFs are positively correlated to Google search volume. Suppose that ASVI is correlated to the institutional attention through channels like fundamental news, then rational institutions would adjust their valuations for the firm following the news and make trading decisions based on the relation between current share price and the institutions' updated firm valuation. In this case, most rational institutions would trade in the same direction, either long or short. Thus, after fundamental news, institutional flows of long and short single-stock ETFs should always be asymmetric in direction, contradicting with the symmetric findings in tables 3 and 4. To further dismiss the alternative hypothesis that my findings above are driven by institutional trading, instead of retail trading, I empirically measure retail trading using the algorithm from Barber et al. (2024), and test for the relation between retail attention on underlying stock and retail trading of single-stock ETFs.

4.3. Retail Trading of Single-Stock ETF and Retail Attention

In this subsection, I directly identify the retail trades in NYSE Trade and Quote (TAQ) data using the algorithm in Barber et al. (2024). Then I investigate how the retail trades of single-stock ETFs are related to the retail attention drawn to the underlying stocks of the ETFs.

Boehmer et al. (2021) propose an algorithm to both measure and sign retail investor trading using NYSE Trade and Quote (TAQ) data. Barber et al. (2024) further improve this algorithm to increase the accuracy rate for the computed signs of retail trades. However, Barber et al. (2024) also experimentally show that both their algorithm and Boehmer et al. (2021) algorithm can correctly identify only one third of all retail trades as retail trades, and this identification rate decreases with bid-ask spreads. Since the bid-ask spreads of single-stock ETFs are large, the identification rate of Barber et al. (2024) algorithm may be even lower than one third for single-stock ETF trades. Therefore, Barber et al. (2024) retail buy and sell volumes computed below only serve as proxies, instead of accurate accounts, for retail trading in each direction. An implication of this caveat is that all retail trading results in this subsection do not provide any information on institutional trading because institutional trades cannot be identified as the complement set of the identified retail trades.

In my sample period, I obtain trade-level data from the NYSE TAQ database for all single-stock ETFs in my sample. I implement Barber et al. (2024) algorithm to identify and sign retail trades. For each single-stock ETF, I compute the daily total dollar volumes of all retail buy trades and all retail sell trades, respectively. It's important to note that Barber et al. (2024) show that their algorithm does not identify all retail trades, and thus these retail dollar volumes are a proxy, instead of an accurate measure, for the intensity of retail trades. Then, I scale the dollar volumes of both retail buy trades and retail sell trades by the market capitalization of the single-stock ETF at the end of the previous trading day. This gives me two dependent variables that measure retail buy activity and retail sell activity for each single-stock ETF on each trading day. This process to compute the dependent variables is summarized in the formulae below:

$$RetailBuy_{i,t} = \frac{\sum_{b \in B_{i,t}} Size_b \times Price_b}{MarketCap_{i,t-1}} \tag{9}$$

$$RetailSell_{i,t} = \frac{\sum_{s \in S_{i,t}} Size_s \times Price_s}{MarketCap_{i,t-1}} \tag{10}$$

where $B_{i,t}$ is the set of all retail buy trades, as identified by Barber et al. (2024), of single-stock ETF i on day t, and $S_{i,t}$ is the set of all retail sell trades, as identified by Barber et al. (2024), of single-stock ETF i on day t. $Size_b$ is the number of shares traded in trade b, and $Price_b$ is the execution price of trade b.

To compare the differences between retail buying and retail selling for single-stock ETF i on day t, I also compute a buy-sell imbalance (BSI) measure similar to that in Barber and Odean (2008), but with dollar volume of retail trades, instead of number of trades:

$$BSI_{i,t} = \frac{\sum_{b \in B_{i,t}} Size_b \times Price_b - \sum_{s \in S_{i,t}} Size_s \times Price_s}{\sum_{b \in B_{i,t}} Size_b \times Price_b + \sum_{s \in S_{i,t}} Size_s \times Price_s} \tag{11}$$

After computing all dependent variables, I run the daily panel regressions below:

$$RetailTradeMeasure_{i,t} = \alpha + \beta_1 \times ASVI_{i,t} + \beta_2 \times Controls_{i,t} + \delta_t + \gamma_i + \epsilon_{it} \tag{12}$$

where $RetailTradeMeasure_{i,t}$ is one of the three dependent variables above for single-stock ETF i on day t. All independent variables are the same as in previous regressions, except that I do not lag independent variables in this regression. This is because the dependent variable in this regression is a direct measure of retail trades. Retail trades are usually both small and not observable by most investors, and thus retail trades are not likely to affect the independent variables, alleviating the reverse causality concern. This concurrent specification is also adopted by Da et al. (2011) in their regression of retail order on retail attention. Standard errors are again double-clustered at ETF and daily level. The result of this regression is presented in table 7.

The first three columns show that retail investors' buying behaviors of long single-stock ETFs are significantly positively related to the retail attention on the underlying stock, while retail buying for short single-stock ETFs is positively but insignificantly related to attention. The next three columns demonstrate a similar pattern for the retail investors' selling behaviors, except that retail selling of short single-stock ETFs is negatively but insignificantly related to retail attention. Therefore, the first six columns show that retail investors both buy and sell more long single-stock ETFs when they are more attentive to the underlying stocks, while the retail buying and selling of short single-stock ETFs do not seem to significantly relate to retail attention.

Comparing the point estimates for coefficients in columns 2 and 5, one can see that the retail buying of long ETFs increases more than retail selling of long ETFs does when retail attention increases. Columns 3 and 6 also suggest that attention-driven retail buying is greater than attention-driven retail selling of short single-stock ETFs. To formally test this proposition, I regress the Buy-Sell-Imbalance (BSI) measure from Barber and Odean (2008) on retail attention and present the results in the last three columns. Columns 7 to 9 confirm that the retail buying increases more than retail selling when retail attention on the underlying stock is higher. Moreover, this increase in buy-sell imbalance is present for both long and short single-stock ETFs, consistent with findings in Barber and Odean (2008).

Combining the findings from all nine columns of table 7, one can get to the following conclusion: for long single-stock ETFs, higher retail attention on the underlying stock leads to increases in both retail buying and retail selling of the ETF. For both long and short single-stock ETFs, the retail buying increases relative to retail selling when retail attention is higher. My findings are consistent with Barber and Odean (2008) story that due to the short-sale constraints of retail investors on single-stock ETFs, attention-driven retail buying is greater than attention-driven retail selling.

In this section, I have shown that the retail investor attention on the underlying stocks is positively related to the flows and turnover ratios of both long and short single-stock ETFs. I further show that the retail buy-sell imbalance of both long and short single-stock ETFs is positively related to retail attention on the underlying stock, consistent with retail attention literature. Although I do not directly measure institutional trading of single-stock ETFs, the fact that retail trading and aggregate trading exhibit similar patterns provides further support to my story that single-stock ETFs are mainly traded by retail investors, instead of institutional investors.

5. Single-Stock ETFs and Stock Price Idiosyncratic Volatility

After showing that single-stock ETFs mainly attract short-term retail investors, I ask the follow-up question of how the introduction of single-stock ETFs affects the stock price. Ben-David et al. (2018) have shown that the inclusion in an ETF's portfolio results in an increase in the stock price volatility. They explain this relation by a theory that ETFs attract short-term oriented liquidity traders, who raise the volatility of ETF prices, and this increase in volatility is propagated to the underlying stock price by the arbitraging activities. Since I have shown that single-stock ETFs mainly attract attention-driven retail investors, the story in Ben-David et al. (2018) is expected to apply to single-stock ETFs as well. To formally test the hypothesis that the launch of single-stock ETFs raises the volatility of underlying stock prices, I conduct an event study with staggered launches of single-stock ETFs to different underlying stocks. In this section, I conduct the staggered event study with two approaches. First, I apply Callaway and Sant'Anna (2021) estimator to estimate the dynamic treatment effect of the introduction of the first long and short single-stock ETFs on an underlying stock. Second, I follow Baker, Larcker, and Wang (2022) and Deshpande and Li (2019) to conduct a difference-in-differences analysis on a stacked dataset to estimate the treatment effect of single-stock ETF launch.

The dependent variable in both analyses is stock-monthly level idiosyncratic volatility, which is defined as the standard deviation of the residuals from the CAPM regression. For each stock-month, I regress the daily excess return of the stock on the daily market risk premium, and then take the root mean square of the daily residuals from this regression as the stock-monthly idiosyncratic volatility. I use the idiosyncratic volatility, instead of the volatility of raw stock return, because if retail investor behaviors are purely driven by attention, rather than risks, their trading would lead to an increase in the volatility that is not explainable by the market risks. By removing the stock price movement component related to market risks, I could more accurately capture the potential impact of attention-driven trading of retail investors. For the main specification, I do not use factor models with more factors, such as the Fama-French 3 factor model, because the treated stocks are mainly growth and large stocks, as shown in table 2, and controlling for growth and size factors may lead to biases. In unreported specifications, I also use idiosyncratic volatility from Fama-French 3 and 5 factor models, and the results are quantitatively weaker but qualitatively similar. For both event study analyses, I compute the monthly idiosyncratic volatility up to March

2025, because the last treatment month in my sample of single-stock ETFs is December 2024. Ideally, I should have more recent idiosyncratic volatility observations, but March 2025 is the most recent data from CRSP as of the time of this writing. For both event study analyses, the stock sample I use is the union of the set of stocks that ever had a tracking single-stock ETF as of December 2024 and the set of the largest 500 US stocks by market capitalization on the first trading day of 2022. I use the never-treated largest US stocks as control groups, because the treated firms are mainly large-cap stocks, as shown in table 2.

In the first event study analysis, I apply the estimator from Callaway and Sant'Anna (2021) to estimate the dynamic treatment effect of the staggered launch of first long or short single-stock ETFs on each underlying stock. I use the doubly-robust method of Sant'Anna and Zhao (2020), as suggested in Callaway and Sant'Anna (2021). I conduct the event study on the launch of the first long single-stock ETF and the first short single-stock ETF respectively. Although in some cases, the first long single-stock ETF and the first short single-stock ETF on a stock are introduced in the same month, there is significant variation between these two treatment events. In my sample, each underlying stock remains treated once it's treated.

The resulting event study plot from the Callaway and Sant'Anna (2021) estimator is presented in figure 4. The first thing we can observe from both panels in figure 4 is that there is no significant pre-trend before the treatments of both long and short single-stock ETFs. The top panel shows an increase in idiosyncratic volatility starting from the second month after the launch of the first long ETF, but this increase is not significant at the 5% level, and this increase reverts to 0 in the eighth month after the treatment. On the other hand, the bottom panel illustrates a larger and longer-lasting increase in the stock idiosyncratic volatility following the treatment of the first short single-stock ETF launch.

Besides the staggered event study with Callaway and Sant'Anna (2021) approach, I also conduct a stacked difference-in-differences analysis as in Baker et al. (2022) and Deshpande and Li (2019) to arrive at an estimate on the static treatment effect, because the size of the effect on idiosyncratic volatility seems to be permanent in figure 4, at least for short single-stock ETFs. For long and short single-stock ETFs respectively, I create one stacked dataset by stacking all sub-experiments together. Each sub-experiment is defined by the treatment month, in which the first single-stock ETF starts trading. For each sub-experiment, I include the monthly observations, from 3 months before treatment to 3 months after treatment, for all never-treated firms and all firms treated in this specific sub-experiment. I choose 3-month windows before and after the treatment month for the stacked difference-in-differences analysis mainly because of data availability issues. There are many treatment events in the last quarter of 2024, whereas the stock price data is updated to the first quarter of 2025 as of the time of this writing. The following regression is estimated on each stacked dataset:

$$\begin{aligned} IdiosyncraticVolatility_{i,t,j} = \alpha + \beta_1 \times Post_{t,j} + \beta_2 \cdot Post_{t,j} \times Treated_{i,j} \\ + \beta_3 \times Controls_{i,t} + \delta_t + \gamma_i + \epsilon_{i,t,j} \end{aligned} \tag{13}$$

where $Treated_{i,j}$ is a dummy equal to 1 if stock i is treated in sub-experiment j. $Post_{t,j}$ is the post dummy equal to 1 if month t is after the treatment month of sub-experiment j, and equal to 0 if month t is before the treatment month of sub-experiment j. I exclude the monthly observations for the treatment month of each sub-experiment from the regression, because the exact treatment day in the month may vary for each treated stock. I include stock-monthly level control variables that may be related to stock volatility, including stock return, Amihud illiquidity measure, and the log of stock market capitalization at the end of the month. I also control for stock and month fixed effects. Because for each sub-experiment, I only include never-treated stocks and stocks that are treated in this sub-experiment, a treated dummy would have been subsumed by the stock fixed effect. The standard errors are double-clustered at the stock and month level, because treatment assignment is at the stock level.

The results of this stacked difference-in-differences analysis are presented in table 8. Similar to the event study plots in figure 4, table 8 shows that the treatment effect of the first short single-stock ETF on stock idiosyncratic volatility is significantly positive, whereas the treatment effect of the long single-stock ETF is positive but insignificant. One can also observe that stock market cap is negatively related to the impact of treatment on the stock idiosyncratic volatility, likely because retail trading is too small to significantly move the prices of large stocks.

In this section, I conduct event study analyses to investigate the treatment effect of the launch of the first long or short single-stock ETF on the idiosyncratic volatility of the underlying stocks. Both the event study plot from Callaway and Sant'Anna (2021) estimator and the stacked difference-in-differences analysis as in Baker et al. (2022) and Deshpande and Li (2019) illustrate similar patterns: stock idiosyncratic volatility increases after the launch of the first short single-stock ETF, whereas the launch of the first long single-stock ETF has an insignificantly positive effect. In general, these results are consistent with the story in Ben-David et al. (2018) that ETFs attract a new group of short-term liquidity traders, whose trading contributes to higher volatility in the underlying security price.

6. Conclusion

This paper is one of the first to investigate the novel instrument of single-stock ETFs. US single-stock ETFs have enjoyed rapid growth in size and popularity since their recent advent. I introduce the underlying mechanisms of single-stock ETFs and reveal the fundamental characteristics of this instrument. I show that single-stock ETFs are primarily held by individual investors and that investors only hold single-stock ETFs for an average of 2 days. Moreover, single-stock ETFs charge 1.13% annually on average, which is much higher than index-based ETFs. I propose to explain the rapid growth of single-stock ETFs' and their popularity among retail investors, despite their hefty fees, by arguing that single-stock ETFs satisfy retail investors' demands for taking short-term leveraged long and short positions on popular stocks. These demands were previously unmet or very expensive for retail investors.

To support this hypothesis, I show that the flows and turnover ratios of both long and short single-stock ETFs are positively correlated with retail attention on the underlying stocks. I also find that the buy-sell imbalance in retail trades of single-stock ETFs increases with retail attention. These findings are consistent with the investor attention theory in the literature (Barber and Odean, 2008). The symmetric relations between retail attention and retail trading in both long and short single-stock ETFs suggest that single-stock ETFs, especially short ETFs, relax retail investors' trading constraints on underlying stocks by allowing them to short-sell and take leverage. Thus, single-stock ETFs serve the purpose of leveling the playing field between retail and institutional investors. The fact that ETF overall flows and turnover ratios exhibit a similar pattern to the retail trades of ETFs provides further support for my theory that single-stock ETFs are mainly traded by retail investors to meet their demands for leverage and short-selling.

I further investigate how the introduction of single-stock ETFs affects the idiosyncratic volatility of the underlying stocks. I show that the idiosyncratic volatility of underlying stock prices increases following the launch of the first short single-stock ETF tracking this stock. This increase lasts for at least 12 months after the first short single-stock ETF starts trading. Following the launch of the first long single-stock ETF, the increase in idiosyncratic volatility of the underlying stock is positive but insignificant, and it reverts to 0 within 8 months after the treatment.

Although the sizes of underlying stocks are much larger than the sizes of single-stock ETFs as of December 2024, the impact of single-stock ETFs could continue to grow in the future, as the single-stock ETF industry is expected to grow rapidly and start to cover smaller and less liquid stocks. As documented in this paper, the daily aggregate trading volume of single-stock ETFs has even reached 20% of the daily trading volume of the underlying stock in some daily observations for certain stocks. Therefore, understanding and continuously monitoring the impact of single-stock ETFs on the financial market could be an important task for both academics and policymakers. This paper serves as the starting point for the research on single-stock ETFs, and more research will be needed to better understand this novel instrument.

Appendix A. Figures

Fig. 1. Daily Standardized ASVI for All Underlying Stocks

This figure plots daily standardized abnormal SVI values for all 24 underlying stocks of single-stock ETFs. The sample period is from Jan 1st, 2020 to Dec 31st, 2024. The ticker symbols of the stocks are given above each subfigure. The series for META is obtained by combining the values for FB before its change in ticker and the values for META after its change in ticker.

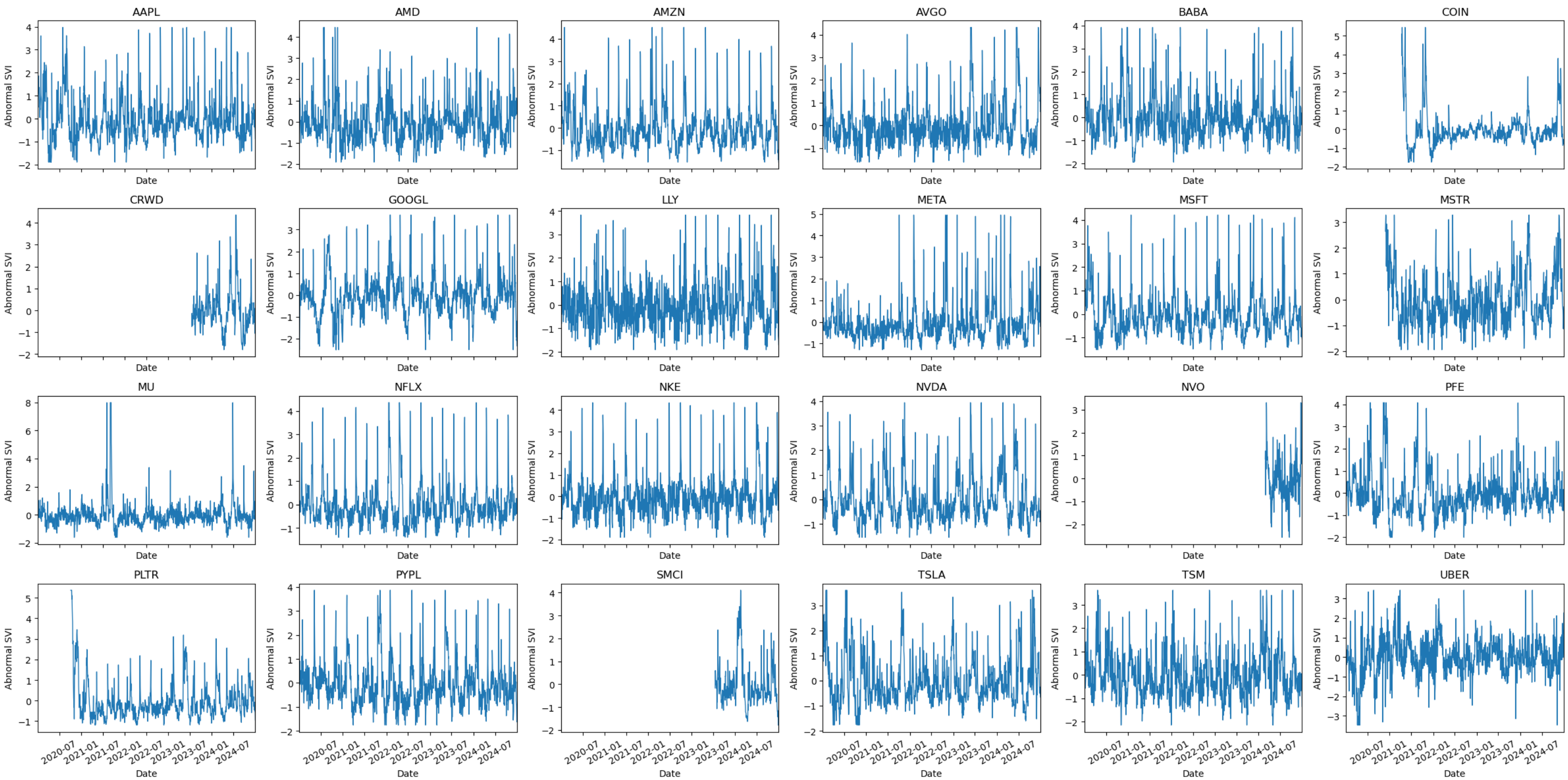


Fig. 2. Growing Size of Single-Stock ETF Industry

This figure plots the time series of daily total market capitalization for all single-stock ETFs in aggregate. The sample period is from July 13th, 2022 to Dec 31st, 2024. The blue line is the daily aggregate total market capitalization for all single-stock ETFs trading in the market, including both long and short ETFs. The orange line is the daily aggregate total market capitalization for all long single-stock ETFs. The green line is the daily aggregate total market capitalization for all short single-stock ETFs. The aggregate size of all single-stock ETFs (blue line) is the sum of the aggregate sizes of all long single-stock ETFs (orange line) and all short single-stock ETFs (green line). Y-axis is the total market capitalization of single-stock ETFs in billions of US dollars.

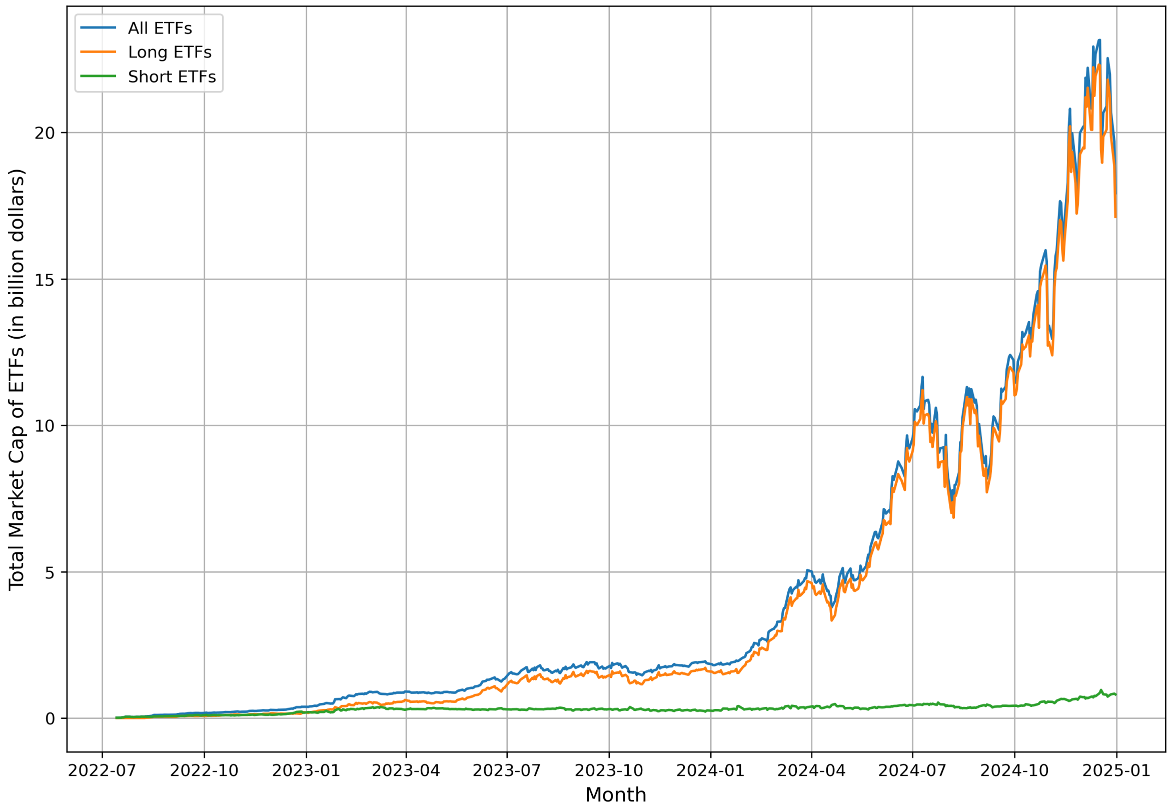


Fig. 3. Institutional Ownership of Single-Stock ETFs

This figure plots the times series of total market capitalization and percentage of single-stock ETFs held by 13-F-reporting institutional investors at a quarterly frequency. The quarterly data are recorded on the last trading day in each calendar quarter. The solid blue line is the aggregate institution-owned total market capitalization of all long single-stock ETFs. The solid orange line is the aggregate institution-owned total market capitalization of all short single-stock ETFs. The aggregate total market capitalizations correspond to the left axis. The dashed blue line is the percentage institutional ownership for long single-stock ETFs in aggregate, computed as aggregate institution-owned total market capitalization for all long single-stock ETFs divided by the sum of total market capitalization for all long single-stock ETFs. The dashed orange line is the percentage institutional ownership for short single-stock ETFs in aggregate, computed as aggregate institution-owned total market capitalization for all short single-stock ETFs divided by the sum of total market capitalization for all short single-stock ETFs. Percentage ownership corresponds to the right axis.

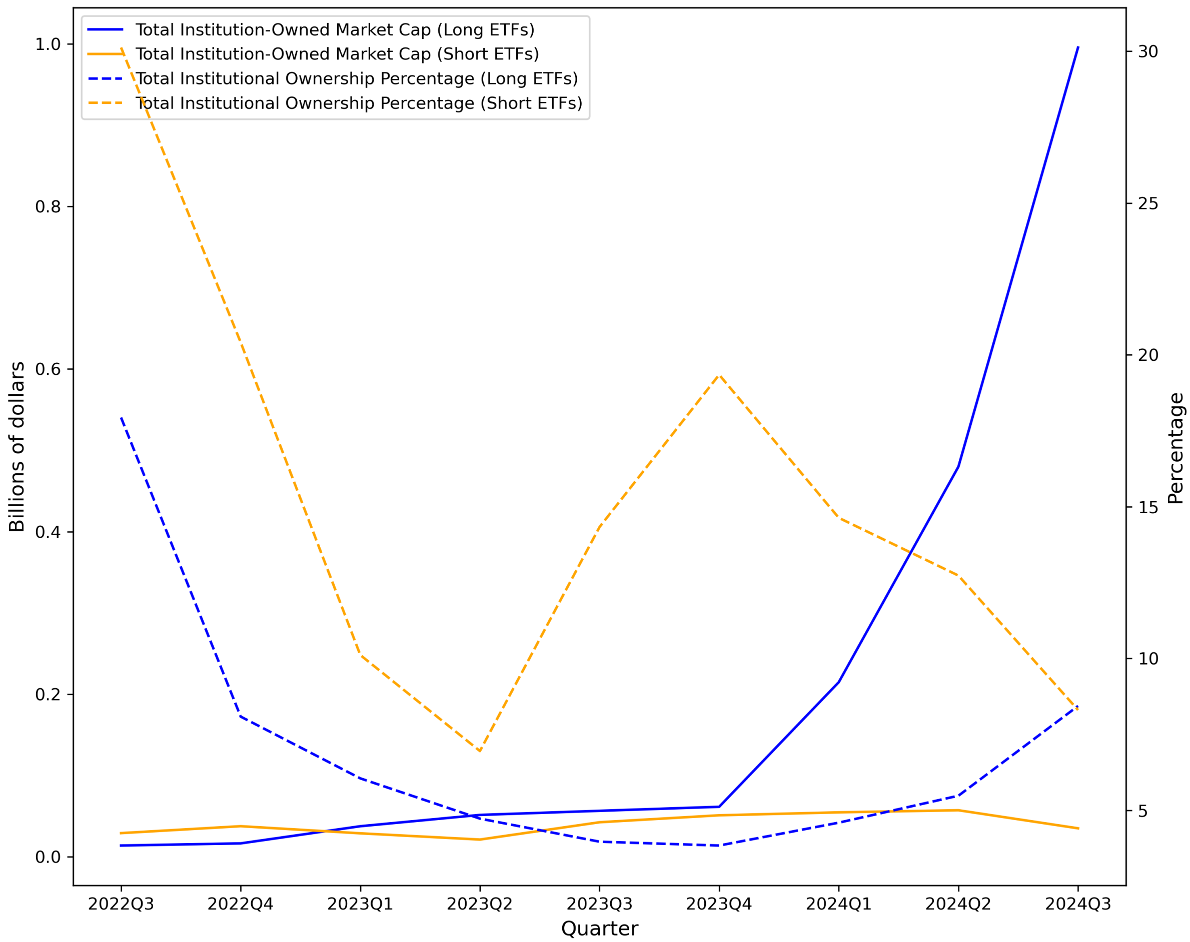
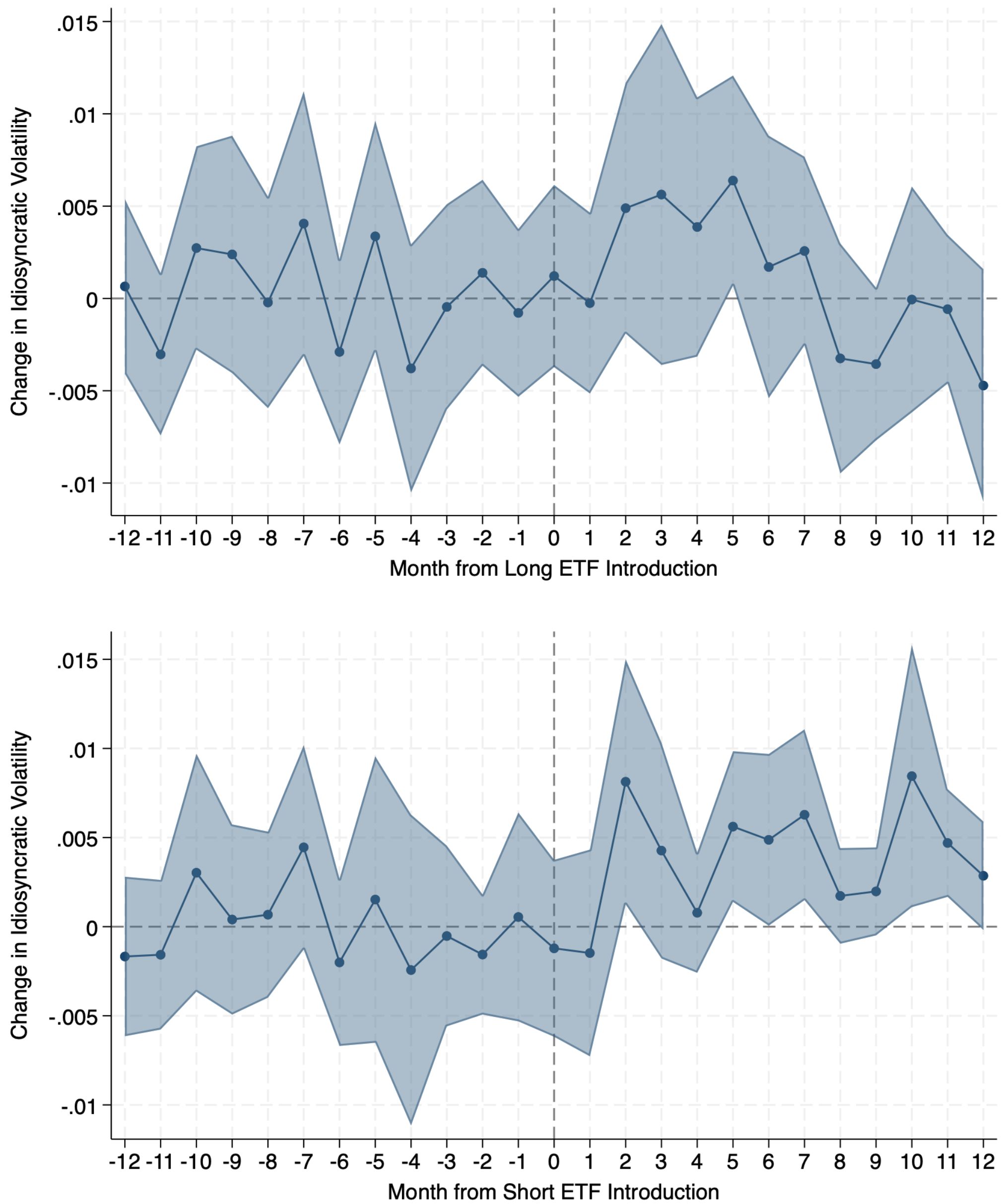


Fig. 4. Stock Idiosyncratic Volatility Around the Launch of First ETF

This figure presents the event study plots on the monthly stock idiosyncratic volatility around the launch of the first long and short single-stock ETF on the stock respectively. Monthly stock idiosyncratic volatility is defined as the standard deviation of the residuals from CAPM regression. y-axis is the regression estimate for the coefficients for each month around the treatment event. x-axis is the number of months from the treatment month, when the first long or short single-stock ETF on the stock starts trading. The top(bottom) panel presents the results from the staggered event study, where the treatment is defined as the launch of the first long(short) single-stock ETF. The event study is conducted as in Callaway and Sant'Anna (2021), using the doubly-robust method from Sant'Anna and Zhao (2020). 95% confidence intervals for the estimates are shaded in blue.



Appendix B. Tables

Table 1: Summary Statistics for Single-Stock ETFs

This table exhibits the summary statistics for the main variables on single-stock ETFs. The sample period is from July 13th, 2022 to Dec 31st, 2024. The first column of numbers is for all long single-stock ETFs. The second column is for all short single-stock ETFs. The last column is data for all single-stock ETFs. The first six rows are variables at ETF-daily level. The seventh to ninth rows are ETF-quarterly level data on institutional ownership of single-stock ETFs. The remaining six rows are ETF-level data. The expense ratio and management fee data are cross-sectional as of June 2024. The first and the last rows are the numbers of observations at ETF-daily level and ETF level respectively. For all other rows, the numbers in front of the parenthesis are the average values, and the numbers in the parenthesis are standard deviations.

	Bet Direction		
	Long	Short	Total
N	11,067 (60.0%)	7,378 (40.0%)	18,445 (100.0%)
Market capitalization (million dollars)	220.64 (707.35)	26.44 (35.40)	142.96 (556.55)
Turnover (percentage)	25.98 (50.43)	83.11 (552.11)	48.83 (352.46)
Proportional flow (decimals)	0.02 (0.17)	0.03 (1.32)	0.02 (0.85)
Dollar flow (thousands)	891.94 (24085.21)	185.85 (4904.16)	609.36 (18912.16)
Premium (percentage)	0.02 (0.26)	0.00 (0.23)	0.01 (0.25)
Quarterly institutional ownership (percentage)	12.63 (18.75)	24.28 (22.31)	17.20 (20.97)
Quarterly individual-owned total market cap (million dollars)	211.82 (618.42)	22.30 (27.93)	135.95 (488.12)
Quarterly institution-owned total market cap (million dollars)	13.72 (59.28)	3.87 (4.04)	9.78 (46.23)
ETF-level tracking error (1 - R-squared)	0.01 (0.01)	0.01 (0.01)	0.01 (0.01)
ETF-level tracking error (abs(beta))	0.08 (0.13)	0.01 (0.02)	0.06 (0.11)
ETF-level expense ratio (percentage)	1.12 (0.11)	1.15 (0.14)	1.13 (0.13)
ETF-level management fee (percentage)	0.93 (0.15)	0.94 (0.18)	0.93 (0.16)
ETF-level leverage	1.96 (0.15)	1.36 (0.47)	1.75 (0.41)
Number of ETFs	47	25	72

Table 2: Summary Statistics for Underlying Stocks

This table exhibits the summary statistics for the main variables on underlying stocks of the single-stock ETFs. The sample period is from Jan 2nd, 2022 to Dec 31st, 2024. All variables are at underlying stock-daily level. I compute 10th-percentile, mean value, standard deviation, and 90th-percentile across all observations for each variable, except N, and list them in four columns respectively. The first row is the number of underlying stock-daily observations in the sample period, including stock-days when no single-stock ETF is trading on the underlying stock. The second to sixth rows are underlying stock-daily variables on stock characteristics, which are computed for every stock-day in my sample. The seventh to tenth rows are stock-daily variables that compare long(short) single-stock ETF characteristics to underlying stock characteristics, and are only computed for stock-days when there is at least 1 long(short) single-stock ETF trading on the underlying stock. The daily Amihud illiquidity ratio for the stocks is computed as absolute value of stock return in decimal, scaled by trading volume in billion dollars. The single-stock ETF market capitalization and dollar trading volume are summed to underlying stock-daily level, and then expressed as a percentage of stock market capitalization and dollar trading volume.

	p10	Mean	SD	p90
N		29,674		
Stock market cap (billion dollars)	15.44	480.94	727.73	1599.51
Stock return (percentage)	-3.10	0.15	3.26	3.37
Stock turnover ratio (percentage)	0.35	1.98	3.06	4.57
Stock book-to-market ratio	0.03	0.27	0.35	0.63
Stock Amihud illiquidity measure (scaled by volume in billions)	0.00	0.09	0.56	0.06
Long ETF aggregate market cap (as percentage of stock market cap)	0.00	0.00	0.00	0.00
Short ETF aggregate market cap (as percentage of stock market cap)	0.00	0.00	0.00	0.00
Long ETF aggregate trading volume (as percentage of stock trading volume)	0.00	0.84	2.29	1.97
Short ETF aggregate trading volume (as percentage of stock trading volume)	0.00	0.17	0.58	0.54

Table 3: Weekly Single-Stock ETF Flows and Retail Attention

This table presents the results of ETF-weekly panel regressions of single-stock ETF flows on retail attention on the underlying stock. The sample period is from the week of July 13th, 2022 to the week of Dec 31st, 2024. The regression is run on three subsamples: all single-stock ETFs, long single-stock ETFs only, and short single-stock ETFs only. The dependent variable is the proportional flow of single-stock ETFs, computed as in Sirri and Tufano (1998). The main independent variable is lagged weekly abnormal SVI (Google Search Volume Index) on the underlying stock. Control variables include common trading-related investor attention proxies and variables widely known to affect fund flows. Control variables for investor attention are absolute value of stock return, stock volatility, and weekly average of daily turnover ratio of underlying stock. ETF return, ETF size, and ETF age are known to affect future ETF flows. All control variables are lagged by 1 week. ETF fixed effect and week fixed effect are included. Standard errors (in parentheses) are double-clustered at ETF and weekly level. *, **, and *** represent significance at the 10%, 5%, and 1% level respectively.

Dependent Variable:	Weekly Proportional ETF Flow		
	All ETFs	Long ETFs	Short ETFs
Lagged stock weekly standardized ASVI	0.083***	0.067**	0.096**
	(0.021)	(0.027)	(0.041)
Lagged abs(stock weekly return)	0.501	0.235	0.905
	(0.366)	(0.413)	(0.886)
Lagged stock average daily turnover	3.197	-1.196	13.488
	(4.163)	(1.330)	(10.734)
Lagged stock weekly volatility	-5.727	-1.341	-12.194
	(4.847)	(1.297)	(11.793)
Lagged ETF weekly return	-0.152	-0.060	-0.160
	(0.107)	(0.077)	(0.477)
Lagged log(ETF market cap)	-0.290*	-0.194**	-0.774
	(0.162)	(0.074)	(0.557)
Lagged log(ETF age)	0.167	-0.033	0.698
	(0.195)	(0.069)	(0.584)
Constant	4.415**	3.527***	10.193
	(2.063)	(1.114)	(7.043)
Fund FE	YES	YES	YES
Week FE	YES	YES	YES
R-squared	0.15	0.43	0.17
N	3,759	2,251	1,508

Table 4: Daily Single-Stock ETF Flows and Retail Attention

This table presents the results of ETF-daily panel regressions of single-stock ETF flows on retail attention on the underlying stock. The sample period is from July 13th, 2022 to Dec 31st, 2024. The regression is run on three subsamples: all single-stock ETFs, long single-stock ETFs only, and short single-stock ETFs only. The dependent variable is the proportional flow of single-stock ETFs, computed as in Sirri and Tufano (1998). The main independent variable is lagged daily abnormal SVI (Google Search Volume Index) on the underlying stock. Control variables include common trading-related investor attention proxies and variables widely known to affect fund flows. Control variables for investor attention are absolute value of daily stock return and daily turnover ratio of underlying stock. ETF return, ETF size, and ETF age are known to affect future ETF flows. All control variables are also lagged by 1 day. ETF fixed effect and day fixed effect are included. Standard errors (in parentheses) are double-clustered at ETF and daily level. *, **, and *** represent significance at the 10%, 5%, and 1% level respectively.

Dependent Variable:	Daily Proportional ETF Flow		
	All ETFs	Long ETFs	Short ETFs
Lagged stock daily standardized ASVI	0.020**	0.007***	0.036
	(0.009)	(0.002)	(0.023)
Lagged abs(stock daily return)	0.035	0.037	0.646
	(0.207)	(0.120)	(0.488)
Lagged stock daily turnover	-0.312	0.068	-0.639
	(0.437)	(0.082)	(1.359)
Lagged ETF daily return	0.275	0.099**	0.673
	(0.301)	(0.039)	(0.939)
Lagged log(ETF market cap)	-0.058	-0.029***	-0.167
	(0.037)	(0.010)	(0.126)
Lagged log(ETF age)	0.035	-0.007	0.135
	(0.041)	(0.010)	(0.123)
Constant	0.817*	0.529***	2.039
	(0.420)	(0.139)	(1.427)
Fund FE	YES	YES	YES
Day FE	YES	YES	YES
R-squared	0.04	0.12	0.08
N	18,299	10,972	7,327

Table 5: Weekly Single-Stock ETF Turnover Ratios and Retail Attention

This table presents the results of ETF-weekly panel regressions of single-stock ETF turnover ratios on retail attention on the underlying stock. The sample period is from the week of July 13th, 2022 to the week of Dec 31st, 2024. The regression is run on three subsamples: all single-stock ETFs, long single-stock ETFs only, and short single-stock ETFs only. The dependent variable is the weekly average of daily turnover ratio of single-stock ETFs. The main independent variable is lagged weekly abnormal SVI (Google Search Volume Index) on the underlying stock. Control variables include common trading-related investor attention proxies and variables known to affect fund trading. Control variables for investor attention are absolute value of stock return and stock volatility. ETF return, ETF size, and ETF age are known to affect future ETF trading. All control variables are lagged by 1 week. ETF fixed effect and week fixed effect are included. Standard errors (in parentheses) are double-clustered at ETF and weekly level. *, **, and *** represent significance at the 10%, 5%, and 1% level respectively.

Dependent Variable:	Weekly ETF Turnover Ratio		
	All ETFs	Long ETFs	Short ETFs
Lagged stock weekly standardized ASVI	0.003***	0.003***	0.002***
	(0.001)	(0.001)	(0.001)
Lagged abs(stock weekly return)	0.037**	0.028	0.023
	(0.017)	(0.017)	(0.015)
Lagged stock weekly volatility	0.145**	0.124*	0.141
	(0.068)	(0.064)	(0.084)
Lagged ETF weekly return	-0.001	0.010**	-0.031
	(0.004)	(0.004)	(0.019)
Lagged log(ETF market cap)	-0.001	-0.001	0.000
	(0.001)	(0.001)	(0.001)
Lagged log(ETF age)	0.002	0.002	0.003
	(0.001)	(0.002)	(0.002)
Constant	0.018	0.029	0.001
	(0.012)	(0.020)	(0.015)
Fund FE	YES	YES	YES
Week FE	YES	YES	YES
R-squared	0.81	0.81	0.84
N	3,759	2,251	1,508

Table 6: Daily Single-Stock ETF Turnover Ratios and Retail Attention

This table presents the results of ETF-daily panel regressions of single-stock ETF turnover ratios on retail attention on the underlying stock. The sample period is from July 13th, 2022 to Dec 31st, 2024. The regression is run on three subsamples: all single-stock ETFs, long single-stock ETFs only, and short single-stock ETFs only. The dependent variable is the daily turnover ratio of single-stock ETFs. The main independent variable is lagged abnormal SVI (Google Search Volume Index) on the underlying stock. Control variables include common trading-related investor attention proxies and variables known to affect fund trading. Control variables for investor attention are absolute value of stock return and daily turnover ratio of underlying stock. ETF return, ETF size, and ETF age are known to affect future ETF trading. All control variables are also lagged by 1 day. ETF fixed effect and day fixed effect are included. Standard errors (in parentheses) are double-clustered at ETF and daily level. *, **, and *** represent significance at the 10%, 5%, and 1% level respectively.

Dependent Variable:	ETF Daily Turnover Ratio		
	All ETFs	Long ETFs	Short ETFs
Lagged stock daily standardized ASVI	0.197**	0.073***	0.323*
	(0.090)	(0.009)	(0.157)
Lagged abs(stock daily return)	4.036	1.046***	8.356
	(2.624)	(0.362)	(7.373)
Lagged ETF daily return	-0.806	0.358***	-3.182
	(1.045)	(0.101)	(3.280)
Lagged log(ETF market cap)	-0.469	-0.142**	-1.373
	(0.353)	(0.059)	(1.249)
Lagged log(ETF age)	0.465	0.011	1.406
	(0.382)	(0.046)	(1.176)
Constant	5.866	2.580***	15.660
	(3.939)	(0.833)	(14.276)
Fund FE	YES	YES	YES
Day FE	YES	YES	YES
R-squared	0.30	0.24	0.34
N	18,299	10,972	7,327

Table 7: Daily Retail Trades of Single-Stock ETFs and Retail Attention

This table presents the results of ETF-daily panel regression of retail trades of single-stock ETFs on retail attention on the underlying stock. The sample period is from July 13th, 2022 to Dec 31st, 2024. There are three dependent variables in the regressions: daily total dollar volume of all retail buy trades, daily total dollar volume of all retail sell trades, and buy-sell imbalance (BSI) measure from Barber and Odean (2008). The first two dependent variables are both scaled by the market cap of the ETF at the end of previous trading day. All retail buy trades and retail sell trades are identified by the algorithm introduced in Barber et al. (2024). For each of the dependent variables, I run the regression using daily observations of all ETFs, long ETFs, and short ETFs respectively. The main independent variable is concurrent abnormal SVI (Google Search Volume Index) on the underlying stock. Control variables include common trading-related investor attention proxies and variables known to affect fund trading. Control variables for investor attention are absolute value of stock return and daily turnover ratio of underlying stock. ETF return, ETF size, and ETF age are known to affect future ETF trading. All control variables are concurrent. ETF fixed effect and day fixed effect are included. Standard errors (in parentheses) are double-clustered at ETF and daily level. *, **, and *** represent significance at the 10%, 5%, and 1% level respectively.

Dependent Variable:	Daily Retail Buy			Daily Retail Sell			BSI		
ETFs:	All	Long	Short	All	Long	Short	All	Long	Short
Stock daily standardized ASVI	0.015**	0.009***	0.004	0.011*	0.006***	-0.001	0.011***	0.008*	0.018***
	(0.007)	(0.001)	(0.011)	(0.006)	(0.001)	(0.010)	(0.003)	(0.004)	(0.006)
Abs(stock daily return)	-0.799	0.172***	-1.683	-0.994	0.140***	-1.976	0.568***	0.565***	0.209
	(1.087)	(0.055)	(2.003)	(1.227)	(0.041)	(2.279)	(0.148)	(0.173)	(0.301)
Stock daily turnover	3.937	0.593***	12.470	4.088	0.621***	12.806	-0.185	-0.007	-0.668
	(3.021)	(0.184)	(8.043)	(3.144)	(0.185)	(8.329)	(0.186)	(0.203)	(0.402)
ETF daily return	0.231	0.041**	0.918	0.296	0.022*	1.183	0.155*	0.079	-0.599***
	(0.232)	(0.018)	(0.885)	(0.288)	(0.013)	(0.968)	(0.082)	(0.110)	(0.137)
log(ETF market cap)	-0.044	-0.010***	-0.142	-0.030	-0.005**	-0.099	-0.000	-0.003	0.004
	(0.037)	(0.004)	(0.139)	(0.027)	(0.002)	(0.101)	(0.004)	(0.006)	(0.014)
log(ETF age)	0.040	-0.008**	0.132	0.032	-0.006**	0.100	-0.068***	-0.081***	-0.037**
	(0.039)	(0.003)	(0.111)	(0.029)	(0.003)	(0.081)	(0.011)	(0.015)	(0.017)
Constant	0.520	0.216***	1.511	0.326	0.117***	0.975	0.388***	0.502***	0.170
	(0.397)	(0.056)	(1.618)	(0.266)	(0.035)	(1.148)	(0.067)	(0.071)	(0.192)
Fund FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
Day FE	YES	YES	YES	YES	YES	YES	YES	YES	YES
R-squared	0.32	0.33	0.39	0.32	0.37	0.39	0.07	0.10	0.13
N	18,252	10,974	7,278	18,252	10,974	7,278	18,001	10,861	7,137

Table 8: Monthly Stock Idiosyncratic Volatility Around First ETF Launch

This table presents the results from stacked difference-in-difference regressions on the monthly stock idiosyncratic volatility around the treatment event of the first single-stock ETF launch on the stock. The treatment events include the launch of the first long and short single-stock ETFs on each underlying stock. The stocks in treated group are the underlying stocks ever tracked by any single-stock ETF before Dec 31st, 2024. The stocks in control group are the 500 largest US stocks by market capitalization as of Jan 2nd, 2022 that are not in treated group. For each sub-experiment, 3 monthly observations before treatment month and 3 monthly observations after treatment month are included. The dependent variable is the stock monthly idiosyncratic volatility, defined as the standard deviation of residuals from daily CAPM model regression. Sample period for monthly idiosyncratic volatility is from April 2022 to March 2025. Post dummy is equal to 1 if the stock-monthly observation is after the treatment month in the respective subexperiment, and equal to 0 if it's before the treatment month. Stock fixed effect and month fixed effect are included. Standard errors (in parentheses) are double-clustered at stock and monthly level. *, **, and *** represent significance at the 10%, 5%, and 1% level respectively.

Dependent Variable:	Idiosyncratic Volatility			
ETF Direction:	Long		Short	
Post dummy	-0.000	-0.000	-0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)
Treated dummy × Post dummy	0.002	0.002	0.003***	0.003***
	(0.002)	(0.002)	(0.001)	(0.001)
Stock monthly return		0.001		0.002
		(0.005)		(0.005)
Stock average daily Amihud illiquidity		-0.001		-0.000
		(0.001)		(0.001)
Stock log(market cap)		-0.005***		-0.004**
		(0.001)		(0.001)
Constant	0.017***	0.067***	0.017***	0.055***
	(0.000)	(0.015)	(0.000)	(0.014)
Stock FE	YES	YES	YES	YES
Month FE	YES	YES	YES	YES
R-squared	0.50	0.51	0.50	0.50
N	56,274	56,274	51,238	51,238

Appendix C. List of US Single-Stock ETFs

ETF Name	Underlying Stock	Direction	Leverage
GraniteShares 2x Long AAPL Daily ETF	AAPL	Long	2
T-Rex 2X Long Apple Daily Target ETF	AAPL	Long	2
Direxion Daily AAPL Bull 2X Shares	AAPL	Long	2
Direxion Daily AAPL Bear 1X Shares	AAPL	Short	1
GraniteShares 2x Long AMD Daily ETF	AMD	Long	2
GraniteShares 1x Short AMD Daily ETF	AMD	Short	1
Direxion Daily AMZN Bull 2X Shares	AMZN	Long	2
GraniteShares 2x Long AMZN Daily ETF	AMZN	Long	2
Direxion Daily AMZN Bear 1X Shares	AMZN	Short	1
Defiance Daily Target 2X Long AVGO ETF	AVGO	Long	2
Direxion Daily AVGO Bull 2X Shares	AVGO	Long	2
Direxion Daily AVGO Bear 1X Shares	AVGO	Short	1
GraniteShares 2x Long BABA Daily ETF	BABA	Long	2
GraniteShares 2x Long COIN Daily ETF	COIN	Long	2
GraniteShares 1x Short COIN Daily ETF	COIN	Short	1
GraniteShares 2x Long CRWD Daily ETF	CRWD	Long	2
T-Rex 2X Long Alphabet Daily Target ETF	GOOGL	Long	2
Direxion Daily GOOGL Bull 2X Shares	GOOGL	Long	2
Direxion Daily GOOGL Bear 1X Shares	GOOGL	Short	1
Defiance Daily Target 2X Long LLY ETF	LLY	Long	2
GraniteShares 2x Long META Daily ETF	META	Long	2
Direxion Daily META Bull 2X Shares	META	Long	2
Direxion Daily META Bear 1X Shares	META	Short	1
T-Rex 2X Long Microsoft Daily Target ETF	MSFT	Long	2
GraniteShares 2x Long MSFT Daily ETF	MSFT	Long	2
Direxion Daily MSFT Bull 2X Shares	MSFT	Long	2
Direxion Daily MSFT Bear 1X Shares	MSFT	Short	1
Defiance Daily Target 2X Long MSTR ETF	MSTR	Long	2
T-Rex 2X Long MSTR Daily Target ETF	MSTR	Long	2
T-Rex 2X Inverse MSTR Daily Target ETF	MSTR	Short	2
Defiance Daily Target 2X Short MSTR ETF	MSTR	Short	2
GraniteShares 2x Long MU Daily ETF	MU	Long	2
Direxion Daily MU Bull 2X Shares	MU	Long	2
Direxion Daily MU Bear 1X Shares	MU	Short	1
Direxion Daily NFLX Bull 2X Shares	NFLX	Long	2

ETF Name	Underlying Stock	Direction	Leverage
T-Rex 2X Long NFLX Daily Target ETF	NFLX	Long	2
Direxion Daily NFLX Bear 1X Shares	NFLX	Short	1
AXS 2X NKE Bull Daily ETF	NKE	Long	2
AXS 2X NKE Bear Daily ETF	NKE	Short	2
Tradr 1.75X Long NVDA Weekly ETF	NVDA	Long	1.75
Leverage Shares 2X Long NVDA Daily ETF	NVDA	Long	2
Direxion Daily NVDA Bull 2X Shares	NVDA	Long	2
GraniteShares 2x Long NVDA Daily ETF	NVDA	Long	2
T-Rex 2X Long NVIDIA Daily Target ETF	NVDA	Long	2
Direxion Daily NVDA Bear 1X Shares	NVDA	Short	1
Tradr 1.5X Short NVDA Daily ETF	NVDA	Short	1.5
GraniteShares 2x Short NVDA Daily ETF	NVDA	Short	2
T-Rex 2X Inverse NVIDIA Daily Target ETF	NVDA	Short	2
Defiance Daily Target 2X Long NVO ETF	NVO	Long	2
AXS 2X PFE Bull Daily ETF	PFE	Long	2
AXS 2X PFE Bear Daily ETF	PFE	Short	2
Direxion Daily PLTR Bull 2X Shares	PLTR	Long	2
GraniteShares 2x Long PLTR Daily ETF	PLTR	Long	2
Direxion Daily PLTR Bear 1X Shares	PLTR	Short	1
AXS 1.5X PYPL Bull Daily ETF	PYPL	Long	1.5
AXS 1.5X PYPL Bear Daily ETF	PYPL	Short	1.5
Defiance Daily Target 2X Long SMCI ETF	SMCI	Long	2
GraniteShares 2x Long SMCI Daily ETF	SMCI	Long	2
GraniteShares 1.25x Long TSLA Daily ETF	TSLA	Long	1.25
Tradr 1.5X Long TSLA Weekly ETF	TSLA	Long	1.5
Direxion Daily TSLA Bull 2X Shares	TSLA	Long	2
GraniteShares 2x Long TSLA Daily ETF	TSLA	Long	2
Leverage Shares 2X Long TSLA Daily ETF	TSLA	Long	2
T-Rex 2X Long TESLA Daily Target ETF	TSLA	Long	2
Direxion Daily TSLA Bear 1X Shares	TSLA	Short	1
GraniteShares 1x Short TSLA Daily ETF	TSLA	Short	1
Tradr 2X Short TSLA Daily ETF	TSLA	Short	2
T-Rex 2X Inverse TESLA Daily Target ETF	TSLA	Short	2
GraniteShares 2x Long TSM Daily ETF	TSM	Long	2
Direxion Daily TSM Bull 2X Shares	TSM	Long	2
Direxion Daily TSM Bear 1X Shares	TSM	Short	1
GraniteShares 2x Long UBER Daily ETF	UBER	Long	2

References

Baker, A. C., Larcker, D. F., Wang, C. C., 2022. How much should we trust staggered difference-in-differences estimates? Journal of financial economics 144, 370–395.

Barber, B. M., Huang, X., Jorion, P., Odean, T., Schwarz, C., 2024. A (sub)penny for your thoughts: Tracking retail investor activity in taq. The Journal of finance (New York) 79, 2403–2427.

Barber, B. M., Huang, X., Odean, T., Schwarz, C., 2022. Attention-induced trading and returns: Evidence from robinhood users. The Journal of finance (New York) 77, 3141–3190.

Barber, B. M., Odean, T., 2008. All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors. The Review of financial studies 21, 785–818.

Ben-David, I., Franzoni, F., Kim, B., Moussawi, R., 2023. Competition for attention in the etf space. The Review of financial studies 36, 987–1042.

Ben-David, I., Franzoni, F., Moussawi, R., 2018. Do etfs increase volatility? The Journal of finance (New York) 73, 2471–2535.

Bessembinder, H., 2025. Returns to constant leverage strategies: General principles and application to levered single-stock etfs, available at SSRN: https://ssrn.com/abstract=5369417.

Boehmer, E., Jones, C. M., Zhang, X., Zhang, X., 2021. Tracking retail investor activity. The Journal of finance (New York) 76, 2249–2305.

Buss, A., Sundaresan, S., 2023. More risk, more information: How passive ownership can improve informational efficiency. The Review of financial studies 36, 4713–4758.

Callaway, B., Sant'Anna, P. H., 2021. Difference-in-differences with multiple time periods. Journal of econometrics 225, 200–230.

Cebrián, E., Domenech, J., 2024. Addressing google trends inconsistencies. Technological forecasting & social change 202, 123318–.

Cheng, M., Madhavan, A., 2009. The dynamics of leveraged and inverse exchange-traded funds. Journal Of Investment Management 7, 43–62.

Cong, L. W., Huang, S., Xu, D., 2024. The Rise of Factor Investing: "Passive" Security Design and Market Implications. NBER working paper series no. w32016, National Bureau of Economic Research, Cambridge, Mass.

Da, Z., Engelberg, J., Gao, P., 2011. In search of attention. The Journal of finance (New York) 66, 1461–1499.

Da, Z., Shive, S., 2018. Exchange traded funds and asset return correlations. European financial management : the journal of the European Financial Management Association 24, 136–168.

Deshpande, M., Li, Y., 2019. Who is screened out? application costs and the targeting of disability programs. American economic journal. Economic policy 11, 213–248.

Elton, E. J., Gruber, M. J., Busse, J. A., 2004. Are investors rational? choices among index funds. The Journal of finance (New York) 59, 261–288.

Evans, R. B., Moussawi, R., Pagano, M. S., Sedunov, J., 2026. Operational shorting and etf liquidity provision. Journal of Financial Economics 180, 104241.

Israeli, D., Lee, C. M. C., Sridharan, S. A., 2017. Is there a dark side to exchange traded funds? an information perspective. Review of accounting studies 22, 1048–1083.

Sant'Anna, P. H., Zhao, J., 2020. Doubly robust difference-in-differences estimators. Journal of econometrics 219, 101–122.

Seasholes, M. S., Wu, G., 2007. Predictable behavior, profits, and attention. Journal of empirical finance 14, 590–610.

Sirri, E. R., Tufano, P., 1998. Costly search and mutual fund flows. The Journal of finance (New York) 53, 1589–1622.

Spiegel, M., Zhang, H., 2013. Mutual fund risk and market share-adjusted fund flows. Journal of financial economics 108, 506–528.

Mutual Fund Manager Effort, Flow-Performance Sensitivity, and Fund Return Persistence

Abstract

This paper documents a novel empirical finding that mutual fund total net assets are an increasing but concave function of past returns. To explain this pattern, I develop a dynamic principal–agent model featuring privately known managerial skill and time-varying managerial effort. Following periods of high fund returns, investors revise their beliefs upward about manager skill but anticipate lower future effort, generating the observed nonlinear relationship between fund size and performance. Consistent with the model's predictions, I provide empirical evidence of both limited return persistence and reduced managerial effort following strong past performance.

1. Introduction

The asset management industry has always attracted much attention from finance researchers because it plays a crucial role in the financial markets. Asset management companies hold almost a third of US corporate equity, almost one fourth of worldwide corporate bonds, and 12% of US treasury and government agency securities (Investment Company Institute Factbook 2022[1]). As of the end of 2021, US asset management companies held 34.6 trillion dollars in financial assets. Among all asset management companies, US mutual funds are the largest group of asset managers, holding 27 trillion dollars in assets. Mutual funds are also closely related to household wealth, as 47.9% of US households own shares in US-registered funds, mostly mutual funds. The median amount of mutual fund assets owned by a mutual fund-owning household is $200,000. Given the close relation between mutual fund performance and household wealth, it is crucial for researchers and policymakers to know whether the mutual fund industry is operating efficiently to maximize investor wealth. In this paper, I present a model in which mutual fund manager effort is dependent on the explicit manager compensation contract. All parties in my model are expected utility maximizers, and investors are Bayesian. My model demonstrates that even if all parties are rational and fund managers are skilled, certain manager compensation structures can lead to low manager effort and consequently low mutual fund returns. This highlights the significant impact of mutual fund manager compensation contract design on investor wealth.

I document a novel empirical finding that the fund's total net assets are an increasing and concave function of past fund abnormal returns. Moreover, I find that year-end fund total net assets are a decreasing and convex function of the portfolio activeness in the following year. Then, I empirically investigate the factors determining portfolio

[1] All statistics about US mutual funds here are accessed from the website https://www.icifactbook.org/pdf/2022_factbook.pdf

activeness, as measured in Amihud and Goyenko (2013), and find that portfolio activeness is decreasing and convex in past fund abnormal returns. This means mutual fund managers adjust the activeness of their portfolios based on their performance in the preceding year. Lastly, I test for mutual fund return persistence and find that the current abnormal return is a decreasing and convex function of past abnormal return plus 1. In other words, fund abnormal return is not persistent, especially for the worst-performing funds. This finding is consistent with Lynch and Musto (2003).

I propose a principal-agent hidden action model to explain all empirical findings presented above. The major innovation of this model is the introduction of a time-varying mutual fund manager effort. This model involves three parties: the mutual fund manager, mutual fund investors, and the mutual fund family. Each party in the model rationally maximizes their own expected utility. Each mutual fund manager possesses a constant skill, which is one component of the expected fund return. The mutual fund manager chooses an effort level, which is costly to them but proportionally increases the expected return of the fund and thus their expected compensation, in order to maximize their expected future utility. The return in the model should be empirically defined as factor model abnormal return because the return attributable to well-known risk factors does not reflect either manager skill or manager effort. The manager effort is not directly verifiable or contractible. Because effort is not directly observable in data, I use portfolio activeness as a proxy for manager effort. The mutual fund investors first observe past fund returns and rationally update their beliefs about manager skill. Then, investors choose the amount of capital allocated to each fund to maximize their expected constant absolute risk aversion utility. It's crucial to notice that my model treats total net assets of the fund, instead of fund flow, as the investor decision variable. The mutual fund family observes manager skill and sets a linear manager compensation contract to incentivize the manager to devote effort. The fund family sets the optimal contract to maximize expected profit.

This model theoretically produces all the empirical findings presented above. In the model presented in this paper, I treat the mutual fund family decision as a constant known by all parties and focus on the decisions made by investors and managers. My model specification assumes that the manager incentive payment is dependent on the magnitude of improvement in current year performance over last year performance. With a higher past return, the manager's expected marginal benefit of effort is lower because it's more difficult to improve over past return, while the expected marginal cost of effort stays the same. Therefore, the managers with higher past returns will devote less effort in the future and tend to have lower future returns. This coincides with my empirical findings of both the decreasing and convex relation between future portfolio activeness and past abnormal return and the lack of persistence in fund abnormal return. My model also predicts that investor capital allocation should be decreasing and convex in future portfolio activeness, consistent with the empirical finding. According to my model, the intuition behind this relation is that an increase in future portfolio activeness brings a greater increase in the variance of future return than in the expected value of future return, resulting in a lower optimal capital allocation for investors with constant absolute risk aversion. In my model, past fund performance affects the optimal investor capital allocation through two channels: investor belief about manager skill and future manager effort. First, a higher past performance will induce Bayesian investors to believe that the manager has a higher skill and to invest more in the fund. Second, given the performance improvement-based compensation structure, a higher past performance will result in a lower future manager effort, which affects both the expected return and the variance of investors' fund investments. The resulting optimal investor capital allocation function from my model is too complex for comparative statics. The numerical simulation of this investor decision function shows that optimal investor capital allocation is an increasing and concave function of past fund abnormal return, exactly matching the empirical finding of the relation between fund total net assets and past abnormal return. I further test the model by empirically computing the numerator and the denominator in the optimal investor capital allocation function for each fund-year. Regression outputs support the optimal investor capital allocation function from the model by confirming that fund total net assets are increasing in the numerator of the optimal investor capital allocation function and decreasing in the denominator.

It's crucial to notice that my model produces fund returns that are not persistent, despite the assumption of constant manager skill. One of the oldest questions in mutual fund literature is whether mutual fund managers have skill, which is usually defined as the ability to persistently generate value for investors (Elton, Gruber, and Blake, 1996; Carhart, 1997; Chevalier and Ellison, 1999). Many previous articles have treated fund return persistence as evidence for the existence of fund manager skill. Berk and Green (2004) decouple manager skill from performance persistence by introducing a rational model in which mutual fund managers have constant skills but do not persistently generate high returns because of investor rationality and diminishing returns to scale. In their model, investors are rational and infer manager skill from observed returns. Because investors devote more capital to funds with higher-skilled managers, managers with higher skills will manage larger funds. Combined with diminishing returns to scale, all funds earn equal abnormal returns in equilibrium. Like Berk and Green (2004), my model in this paper can also generate a lack of persistence in fund returns with constant manager skill assumed. Therefore, my model also supports the notion that mutual fund return persistence and mutual fund manager skill are not closely related concepts. The main difference between my model and Berk and Green (2004) lies in the underlying mechanisms that lead to the lack of return persistence. Berk and Green

(2004) rely on the explicit assumption of diminishing returns to scale to mechanically generate the lack of persistence in fund return. Nevertheless, my model relates future fund return to past fund return through the time-varying manager effort. Moreover, my model is more flexible than Berk and Green (2004) in generating different patterns of return persistence. Although the empirically motivated manager contract leads to a lack of return persistence according to my model, my model could theoretically produce any pattern of return persistence by arbitrarily choosing different manager contract specifications, which are completely possible in reality. This feature highlights the importance of fund manager compensation design in shaping the fund performance. In contrast, the model in Berk and Green (2004) cannot generate persistent returns because of the mechanical relation between fund size and fund return they assume. The flexibility and generalizability of my model is useful as the existence of mutual fund return persistence is still debated in the literature. My model also points to a new direction of separately investigating return persistence for funds with different compensation structures. Time-varying manager effort in my model is an alternative mechanism to the diminishing returns to scale in Berk and Green (2004) for generating the lack of return persistence. Time-varying manager effort might even be the cause of the diminishing returns to scale, but this claim needs further investigation.

My interpretation of the finding that fund total net assets are an increasing and concave function of past fund performance is related to the well-known topic of flow-performance sensitivity (Sirri and Tufano, 1998; Chevalier and Ellison, 1997). In my model, I treat fund total net assets as the decision made by mutual fund investors, whereas the proportional fund flow measure is traditionally regarded as the investor decision variable in flow-performance sensitivity literature. Therefore, one of the innovations of this paper is proposing a new variable for mutual fund investor decision. I argue that using fund total net assets as the investor decision variable is superior to using proportional flow, because total net assets, unlike the proportional flow measure, are not subject to the criticism proposed by Spiegel and Zhang (2013). According to Spiegel and Zhang (2013), the main problem with the traditional proportional flow perspective is that it implicitly links the aggregate flow into the whole mutual fund industry to the return distributions of individual funds with different sizes, which is not supported by empirical analysis. My model has implications for fund total net assets, and it does not necessarily link the aggregate flow into all mutual funds to the return distributions of different-sized funds because the optimal flow into the mutual funds in my model is not dependent on the initial fund sizes, unlike in traditional proportional flow models. Therefore, the criticism by Spiegel and Zhang (2013) against the proportional flow measure does not apply to the fund total net asset measure or my model.

The convex flow-performance sensitivity is generally regarded as a violation of the rational investor proposition, so behavioral explanations have been proposed (Sirri and Tufano, 1998). Nevertheless, I show in this paper that if researchers focus on another investor decision variable, fund total net assets, the empirically observed investor decisions will be consistent with the predictions of rational investor models. Therefore, I argue that the debate around the convex flow-performance sensitivity might just be a result of the inappropriate operationalization of investor decision.

The investor belief part of my model has a similar intuition to Franzoni and Schmalz (2017), in that both models generate variations in investor beliefs about manager ability by allowing variations in the signal-to-noise ratio of past returns. In Franzoni and Schmalz (2017), they find that FPS is higher when the realized market excess return is less extreme. They argue that more extreme market states bring higher noise to the signal about manager ability, and thus the investors' response is less sensitive to the signal when the market states are more extreme. In my model, manager effort serves the function of influencing the signal-to-noise ratio. Because the noise in returns is assumed to be independent of manager effort, whereas the expected return increases with manager effort, a higher manager effort essentially reduces the proportion of noise in the observed past returns, thus increasing the signal-to-noise ratio.

The mutual fund industry is a natural setting for agency problems to arise because the incentives of mutual fund managers and mutual fund investors are not always aligned. In mutual fund literature, researchers usually focus on risk-shifting as a typical result of the principal-agent problem. Risk-shifting is the additional risk exposure fund managers intentionally take on for their own personal benefit, usually at the investors' expense. Risk-shifting also exists in the hedge fund industry (Hodder and Jackwerth, 2007; Aragon and Nanda, 2012). Researchers have investigated the risk-shifting that arises in the mutual fund industry as a consequence of implicit fund flow incentives (Brown, Harlow, and Starks, 1996; Chevalier and Ellison, 1997) and explicit compensation incentives (Lee, Trzcinka, and Venkatesan, 2019). The mutual fund risk-shifting problem is not exactly the same as the standard principal-agent hidden action problem as in Holmström (1979) and Holmström and Milgrom (1991). Nevertheless, nobody has ever applied the standard principal-agent hidden action model to the mutual fund industry as I do in this paper because many argue that this model is not applicable to the setting of mutual funds. This reasoning is explained in Ou-Yang (2003). Ou-Yang (2003) states that in mutual fund industry, a single decision made by a fund manager on portfolio holdings affects both return and risk simultaneously, whereas in standard principal-agent models, agents either control only one factor or can control two factors separately. Therefore, Ou-Yang (2003) argues that standard principal-agent model as in Holmström and Milgrom (1991) is not applicable to mutual fund industry. However, Massa and Patgiri (2009) find that mutual fund managers with higher incentives can achieve higher risk-adjusted returns through dynamic portfolio rebalancing, which is an example of fund manager effort. This empirical finding shows that the mutual fund managers can control risk and return separately by exerting effort, undermining the argument by Ou-Yang (2003). Massa and Patgiri (2009) also find that such superior performance is persistent. Following the findings of Massa and Patgiri (2009), I propose a model that is similar to the standard principal-agent model to describe the interactions among mutual fund managers, investors, and fund families. In this model, I focus on risk-adjusted return and treat effort as a manager's choice that only affects the risk-adjusted return. My current model uses only the explicit compensation structure to determine manager effort, consistent with Lee et al. (2019). The important influence of explicit compensation contracts on manager behaviors is also supported by recent empirical results in Ibert, Kaniel, Van Nieuwerburgh, and Vestman (2018), Ma, Tang, and Gómez (2019), and Evans, Gómez, Ma, and Tang (2024). The implicit manager incentive, as discussed in Brown et al. (1996) and Chevalier and Ellison (1997), could be added to my model as a future extension, and it would be interesting to see the impact of implicit incentive on manager effort in my model. Koijen (2014) presents an example of such a model, in which managers respond to the expectation of future flow and dynamically respond to fund and benchmark returns.

The remainder of the paper is organized as follows. Section 2 describes the data used in the paper and presents the empirical findings. Section 3 discusses the general setup of the model. Section 4 solves the model and relates it to the empirical findings. Section 5 concludes the paper.

2. Data and Empirical Analysis

The monthly and daily factor returns are obtained from Kenneth French's website. I use data from the CRSP Survivor-Bias-Free US Mutual Fund Database from January 1, 2010, to December 31, 2019. I first include all US non-specialty domestic equity funds identified by Lipper objective codes listed in Spiegel and Zhang (2013). Then, following Amihud and Goyenko (2013), I exclude all index funds and funds with less than 70% of assets invested in common stocks. I identify index funds by using the index fund flag data from the CRSP database. Elton et al. (1996) document the survivor bias in some mutual fund databases. They point out that mutual funds with total net assets of less than $15 million are not required to regularly file reports and that these funds backfill all historical performance once they reach the $15 million threshold. Because the funds that have never reached the $15 million threshold will not appear in the database altogether, there would be a survivorship bias if we include beginning observations with less than $15 million in total net assets for any fund. Thus, I drop all beginning observations with total net assets of less than $15 million for each fund. Note that I don't drop observations with total net assets less than $15 million if there is an earlier observation of total net assets greater than $15 million, because this will create gaps in the data and again result in survivorship bias by eliminating poorly performing observations. Evans (2010) suggests eliminating incubation bias by excluding observations on dates earlier than the first offer date recorded in the CRSP database. I follow the suggestion and exclude such observations.

2.1. Determinants of Fund Total Net Assets

Sirri and Tufano (1998) define mutual fund flow as the percentage change in mutual fund total net assets, net of fund return. Since then, the literature has focused on proportional mutual fund flow as the variable of interest that represents investor decisions. However, I argue that investors' total capital allocation, rather than proportional fund flow, is the actual outcome of their decision-making process. As argued by Spiegel and Zhang (2013), using proportional fund flow as the investor decision variable will result in regression misspecification and produce biased estimates for coefficients. The total net asset measure avoids this problem, and thus I argue that researchers should use fund total net assets as the investor decision variable. Intuitively, I suggest that proportional mutual fund flow is a meaningless byproduct, rather than the outcome, of the investor decision-making process. Based on this reasoning, I investigate the factors affecting mutual fund investor decisions by using fund total net assets as the dependent variable.

I first extend the literature on mutual fund flow-performance sensitivity. Sirri and Tufano (1998) find that the proportional mutual fund flow is an increasing and convex function of past fund raw return. They focus on past raw return, which is the most salient value to investors, and explain the pattern with the behavioral argument of search cost. In this paper, I want to focus on the rational behaviors of mutual fund investors; thus, I use the Carhart 4-factor model abnormal return as the proxy for past fund performance. If the investors are Bayesian, they would update their beliefs about manager skill by observing past returns, and the part of past returns that is attributable to well-known risk factors should not affect investors' rational beliefs. This gives the first testable hypothesis.

> *Hypothesis 1:* The total net assets of the fund are increasing in the past fund Carhart 4-factor model abnormal return.

In addition to past fund performance, future portfolio activeness is also related to the fund total net assets. Amihud and Goyenko (2013) find that fund activeness and selectivity are decreasing in fund total net assets. Moreover, both Amihud and Goyenko (2013) and Cremers and Petajisto (2009) show that fund activeness positively influences fund return. From the perspective of a rational investor, the capital allocation to each fund is determined by the fund return and risk characteristics. Hence, because future fund portfolio activeness affects future fund return, a rational investor who correctly expects future portfolio activeness should adjust his capital allocation based on the future portfolio activeness. If fund activeness only positively affects the fund future expected return and does not affect risk exposure, a rational investor would allocate more capital to funds with higher future activeness. Therefore, I will also test this hypothesis in the regression analysis.

> *Hypothesis 2:* Fund total net assets are increasing in future fund portfolio activeness.

Amihud and Goyenko (2013) propose using 1 minus factor model R-squared as an alternative to the "active share" measure (Cremers and Petajisto, 2009) to measure portfolio activeness. It can also be interpreted that a higher factor model R-squared signals a higher degree of closet-indexing by the mutual fund. Amihud and Goyenko (2013) show that their findings are robust to different factor models used as benchmarks. I follow Amihud and Goyenko (2013) and use $(1 - \text{Carhart}R^2)$ as a proxy for portfolio activeness. Following Amihud and Goyenko (2013), I estimate the Carhart 4-factor model by using monthly fund returns from CRSP. Because I want to generate one activeness value for each fund-year, I regress 12 observations of monthly returns on 12 monthly observations of Carhart 4 factor returns for each fund-year. This might seem to be very few observations, but Amihud and Goyenko (2013) also use only 24 monthly observations to estimate the model. The Carhart R-squared values in my estimation, as shown in table 1, are also very close to the values reported in Amihud and Goyenko (2013). I repeat the exercise using daily returns to estimate the Carhart model for each fund-year. Most patterns found using monthly data are preserved when daily data is used. The summary statistics for the final data used in the regression analysis are listed in table 1.

I examine the relationship between the investor decision variable, year-end fund total net assets, and past fund performance and the relationship between total net assets and future portfolio activeness respectively by running the following regressions:

$$TNA_{it} = \gamma_{0,1} + \gamma_{1,1} \cdot (1 + r_{it}) + \gamma_{2,1} \cdot (1 + r_{it})^2 + \text{controls}_{it} + \delta_{i,1} + \epsilon_{it,1} \tag{1}$$

$$TNA_{it} = \gamma_{0,2} + \gamma_{1,2} \cdot e_{i,t+1} + \gamma_{2,2} \cdot e_{i,t+1}^2 + \text{controls}_{it} + \delta_{i,2} + \epsilon_{it,2} \tag{2}$$

where TNA_{it} is the year-end total net asset, in millions of dollars, for fund i in year t, $r_{i,t-1}$ is the abnormal return from the Carhart 4-factor model, and e_{it} is $(1-$Carhart $R^2)$ for fund i in year t. I will use 1 plus the Carhart abnormal return as the independent variable throughout my paper because this is the value that matters according to my model. As will be shown later, the incentive contract and thus manager effort cannot be directly dependent on the Carhart abnormal return, instead of 1 plus the Carhart abnormal return, because optimal manager effort would then become negative with some past return values, yielding unmeaningful interpretations and contradicting the model assumption. I use the same set of control variables as in Amihud and Goyenko (2013), except for the fund total net asset because the fund total net asset is the dependent variable in the regressions above. The expense ratio is the fraction of total investment that shareholders pay for operating expenses, which includes the 12b-1 fee. The annual turnover ratio is defined as the minimum of aggregated sales or aggregated purchases of securities divided by the average twelve-month TNA of the fund. Manager tenure is computed as 1 plus the difference between the observation year and the year the fund manager took over the fund, as recorded by CRSP. Fund age is computed as 1 plus the difference between the observation year and the year of the first offer date. I do not lag the control variables as in Amihud and Goyenko (2013) because the year-end total net asset is an investor decision made at the end of the year, when the control variables from the current year are observable. The outputs of the regressions above are displayed in table 2.

The first important point to notice in table 2 is that the year-end fund total net assets are strongly increasing in past abnormal returns. This result is 1% significant in both regressions containing the past return as an independent variable. Moreover, the concavity of the past return is also 1% significant in both regressions. Thus, empirical analysis suggests that investors' capital allocation is an increasing and concave function of past fund performance, confirming hypothesis 2.1.

The results in table 2 also show that the year-end fund total net asset is a decreasing and convex function of future fund portfolio activeness. In columns 2 and 3, the negative relationship between TNA and future portfolio activeness is significant at the 1% level. In both columns 2 and 3, TNA is a convex function of future portfolio activeness, and this convexity is 5% significant in column 3. Overall, the empirical analysis demonstrates that investors' capital allocation to each fund is decreasing and convex in future fund portfolio activeness. This result suggests that investors invest less in the funds with higher future portfolio activeness, rejecting hypothesis 2.1. This is a curious result because previous studies have shown that portfolio activeness positively predicts fund performance. The result in table 2 suggests that investors invest less despite the higher expected return. There are two possible reasons for this finding. First, investors do not correctly expect future fund portfolio activeness. Second, fund activeness increases risk more than it increases expected return. Since this paper is mainly focused on rational investor behavior, I will investigate the second possible reason later in the paper.

In all columns in table 2, the signs of all control variables remain the same, and the significance of all control variables remains very high.

2.2. *Portfolio Activeness and Past Performance*

Having shown that investor decisions are related to the future portfolio activeness of the fund, I will further investigate what determines future portfolio activeness. Portfolio activeness is a decision made by the portfolio management team, especially the fund manager. As shown in previous literature on the mutual fund agency problem (Brown et al., 1996; Lee et al., 2019), mutual fund managers' decisions are influenced by past fund performance for various reasons. Therefore, I test whether fund portfolio activeness is related to past fund Carhart 4-factor model abnormal returns by running the following panel regression:

$$e_{it} = \gamma_{0,3} + \gamma_{1,3} \cdot (1 + r_{i,t-1}) + \gamma_{2,3} \cdot (1 + r_{i,t-1})^2 + \text{controls}_{i,t-1} + \delta_{i,3} + \epsilon_{it,3} \tag{3}$$

where e_{it} is $(1 - \text{Carhart } R^2)$ for fund i in year t, and $r_{i,t-1}$ is the abnormal return from the Carhart 4-factor model. Following Amihud and Goyenko (2013), I use the same set of control variables and lag all control variables by one year because when managers decide on the portfolio activeness, the control variables for the current year are not observable. The output of the regression is reported in table 3.

As shown in table 3, the current portfolio activeness, proxied by $1 - \text{Carhart } R^2$, is decreasing in the past Carhart abnormal return, significant at the 1% level in all three columns. Column 3 also shows that the portfolio activeness is a convex function of past return at the 1% significance level. Column 1 shows that past Carhart abnormal return alone explains 31.2% of the variation in fund portfolio activeness, demonstrating the high explanatory power of past Carhart abnormal return. Moreover, all lagged control variables have significantly positive impacts on the portfolio activeness at the 1% level.

Among all control variables, portfolio activeness is an increasing but concave function of lagged fund size. Overall, table 3 shows that the negative relationship between current portfolio activeness and past performance is very significant and robust. The convexity of the relation between current portfolio activeness and past fund performance is also very significant.

2.3. Mutual Fund Return Persistence

I will now examine the persistence of mutual fund abnormal returns. There has been mixed evidence for mutual fund return persistence in the literature. I will test whether past fund performance is indicative of future fund performance. Again, I use the annual abnormal return from the Carhart 4-factor model as the measure of fund performance. The equation for the panel regression is given below.

$$r_{it} = \gamma_{0,4} + \gamma_{1,4} \cdot (1 + r_{i,t-1}) + \gamma_{2,4} \cdot (1 + r_{i,t-1})^2 + \text{controls}_{i,t-1} + \delta_{i,4} + \epsilon_{it,4} \tag{4}$$

I again use 1 plus abnormal return as an independent variable to be consistent with my model presented later in the paper. Following Amihud and Goyenko (2013), I include the same six lagged variables that might affect future returns. The results of the above regression are given in table 4.

Results in table 4 show that the future fund abnormal return is a decreasing and convex function of past abnormal return. This implies that there is no persistence in the Carhart abnormal return of mutual funds. This result is robust and is 1% significant in both columns, regardless of whether control variables are included or not. Moreover, column 1, which only includes two independent variables related to past return, has a very high R-squared of 0.907, meaning the past abnormal return explains most of the variation in future abnormal return.

3. General Model Setup

Having presented the empirical findings, I am going to discuss the general setup of a rational model that explains all these findings. The model consists of three parties: the mutual fund manager, the mutual fund investor, and the mutual fund family. Each party rationally maximizes their own expected utility. In this model, every party is assumed to know the others' optimal decision strategy for sure. This assumption could be justified because the three parties repeatedly interact with each other in reality. Therefore, each party could learn about the others' strategies over time, even if they are not perfectly rational. All characteristics of each party, such as the risk aversion of investors and the cost of effort for managers, are also assumed to be known by every party, with the only exception being manager skill, which is not known with certainty by investors. Manager skill is a characteristic of mutual fund managers. It varies cross-sectionally but is constant over time. In other words, I assume there is no learning or networking for mutual fund managers. Manager skill in this sense could be thought of as the intellectual capabilities of a manager (Chevalier and Ellison, 1999), instead of knowledge or industry relations, which tend to change over time. In this model, manager skill is known with certainty to both fund managers themselves and the fund families that employ the managers, but it is not known to fund investors. This assumption is different from some papers, such as Lynch and Musto (2003), Berk and Green (2004), and Dangl, Wu, and Zechner (2008), in which the learning about manager skill is symmetric between fund investors and managers themselves. My assumption that fund managers know their own skill levels is easy to justify. I also assume that the employers know manager skills because employers interact with managers on a daily basis, and thus it is much easier for employers to infer manager skills by observing the managers' daily behaviors than for fund investors. Moreover, as I discussed previously, manager skill in this model could be thought of as a manager's intellectual capability, which, unlike industry relations, could even be assessed through standardized tests (Chevalier and Ellison, 1999). Employers have the power to ask managers to take tests, so it is easy for a fund family to learn about the skill levels of its own fund managers. In contrast to fund families, fund investors have relatively limited information about fund managers. In the model, I assume fund investors rationally learn about the skill level of a fund manager by observing the past performance of the manager. Therefore, fund investors gradually learn about fund manager skill. Their beliefs become more precise over time, but they never know manager skill with certainty. The return of a fund i at time t is assumed to be the following function.

$$r_{it} = s_i \cdot e_{it} + \epsilon_{it} \tag{5}$$

The return of the fund, r_{it}, is assumed to be the product of manager skill, s_i, and manager effort, e_{it}, subject to a normal noise, $\epsilon_{it} \overset{iid}{\sim} N(0, \sigma_\epsilon^2)$. σ_ϵ^2 is a constant that is known to everyone. Manager skill, s_i, is assumed to be normally distributed to allow investors to apply the Kalman filter to update their beliefs about manager skill. Manager effort, e_{it}, is assumed to be non-negative. Note that in this specification, a fund manager is able to increase the expected return of a fund without increasing the volatility. Some have argued that this is not possible (Ou-Yang, 2003). However, Massa and Patgiri (2009) empirically show that higher-incentive managers can willingly increase the return, even after adjusting for risk exposure. The channel through which mutual fund managers increase risk-adjusted returns, as demonstrated in Massa and Patgiri (2009), is very similar to the concept of effort I am using in this model. Therefore, I argue that the form of the fund return structure I am assuming here is realistic. Now, I will discuss the decision-making process for each party.

3.1. Mutual Fund Manager Decision

Mutual fund managers gain utility from compensation, whereas there is also disutility from their efforts. In each period, a manager knows the compensation structure set by his employer and rationally maximizes the expected end-of-period utility by choosing the optimal effort level. The effort level is chosen at the beginning of a period, whereas utility is realized at the end of the period. Following Holmström and Milgrom (1991), I assume that the managers have constant absolute risk aversion and that their utility consists of two parts: end-of-period total wealth and current-period cost of effort. The cost of effort is known to the manager given his choice of effort level, whereas there is uncertainty in the total wealth level because compensation depends on the stochastic fund return. The managers are assumed to focus only on current period utility and don't take any future costs or future benefits into account. Equivalently, I assume there are no career concerns for the managers because career concern, which has been widely discussed in the literature (Chevalier and Ellison, 1999; Ellul, Pagano, and Scognamiglio, 2020), is not the focus of this paper. This assumption simplifies the model. The manager i's optimization problem at the beginning of time t is given as:

$$\begin{aligned} &\max_{e_{it}} \ \mathbb{E}_{t-1}\left[-e^{-\lambda_i \cdot (I_{it} - \gamma_i \cdot s_i \cdot e_{it}^2)}\right] \qquad (6) \\ &\text{s.t.} \quad I_{it} = I_{i,t-1} + \alpha_i + \beta_i \cdot f(r_{it}) \end{aligned}$$

where e_{it} is the time t effort level chosen by manager i at the beginning of time t. λ_i is the risk aversion coefficient for manager i. γ_i is the cost of effort parameter for manager i. Because the cost of effort term is subtracted from the manager wealth when computing the utility, γ_i can be interpreted as a term that converts the cost of effort into monetary value. The cost of effort also depends on the manager skill, s_i. It can be interpreted as the opportunity cost of effort, because managers with higher skills can usually earn more money elsewhere had they not chosen to devote effort to portfolio management activities. Including s_i in the manager's cost of effort also brings the convenient benefit of eliminating s_i from the manager's optimal effort function. If the optimal manager effort is a function of s_i, then the return, $r_{it} = s_i \cdot e_{it} + \epsilon_{it}$, will become a function of s_i^2. Because in this case, the signal, r_{it}, is not a linear transformation of the investors' variable of interest, s_i, investors would not be able to apply the Kalman filter to update their beliefs about manager skill, making the investor decision problem impossible to solve recursively. As in standard hidden-action models, the cost of effort is assumed to be a quadratic function of effort, e_{it}, to make the manager decision problem well-behaved. I_{it} denotes the wealth level of manager i at the end of period t. This wealth level is the sum of the wealth level in the last period, $I_{i,t-1}$, and the compensation given to the manager in the current period, t. Following Holmström and Milgrom (1991), I assume a linear contract that comprises a fixed lump-sum payment, α_i, and an incentive payment that depends on the measured performance, which is the fund return, r_{it}, in my case. In later empirical analysis, I would further define r_{it} to be the abnormal return from the 4-factor model Carhart (1997), because fund families should not reward managers for returns that could be attributed to well-known risk factors. β_i measures the intensity of the performance-based incentive payment. Both α_i and β_i are determined by the fund family as solutions to the fund family's optimization problem, and both parameters are known to both fund manager and fund investors. It's worth noting that I assume the incentive payment depends on fund return, rather than fund size or fund revenue, which is usually a fixed percentage of fund size (Elton, Gruber, and Blake, 2003). This compensation structure is chosen to simplify the manager decision problem. If I include fund size or fund revenue in the manager incentive payment, the manager will consider both current-period return and next-period fund size when deciding on current-period effort, because current fund performance only affects fund flows in the future. This would introduce the expectation of the investor fund flow function into the manager's optimization problem, making the problem very complex, as will be shown later in this paper. It's a future direction to extend the model, and it would be very interesting to see the implications of a revenue-based incentive payment structure through the lens of this model. However, for the current version, I would only consider a performance-based incentive payment.

Note that there are many possible specifications for $\beta_i \cdot f(r_{it})$, and different specifications would produce different results regarding managers' optimal effort. Because I do not have data on the mutual fund manager incentive contracts, I do not know the exact functional form of $f(r_{it})$ in reality. Nevertheless, in my model, the mutual fund manager incentive contract determines the mutual fund manager's decision on effort. Therefore, by empirically finding the factors that affect future manager effort, I can gain insights into the manager incentive contract. Because manager effort is not empirically observable, I will use the portfolio activeness measure as a proxy for mutual fund manager effort. It's helpful that $(1 - \text{Carhart } R^2)$ is non-negative, consistent with the model assumption about manager effort. Nevertheless, this is not a perfect measure of manager effort because, while closet-indexers generally tend to have a high Carhart R-squared, a high Carhart R-squared does not necessarily mean that the manager is not devoting much effort to managing the portfolio. Moreover, closet-indexing is not the only way for fund managers to avoid devoting effort. For example, if a manager randomly throws a die to determine the portfolio holdings, the manager is not putting in much effort, whereas the Carhart R-squared is likely to be low. It's also worth pointing out that the practice of using a proxy for manager effort itself is somewhat contradictory to the hidden-action idea of the model. Carhart R-squared can be easily computed by investors and fund families, so if Carhart R-squared indeed measures manager effort, a hidden-action model would not be appropriate for modeling the situation. In fact, I think it might be more realistic to think of Carhart R-squared as a contractible noisy signal for the unobservable manager effort. In this case, the appropriate model would be a hidden-action model with a contractible noisy signal. This would be another way to improve the model in the future, but I am not going to pursue this model at the current stage because, again, it would introduce additional complexity to the model. My empirical results in table 3 show that future portfolio activeness is a decreasing and convex function of past fund abnormal return. Based on this finding, I propose to set $f(r_{it}) = \dfrac{1+r_{it}}{1+r_{i,t-1}}$. Therefore, the complete mutual fund manager payment becomes $\alpha_i + \beta_i \cdot \dfrac{1+r_{it}}{1+r_{i,t-1}}$, which includes a fixed payment and an incentive payment that is dependent on the magnitude of improvement in the current year return over the last year return. The optimal manager effort in response to this contract specification would be a decreasing and convex function of last year return, exactly matching the empirical finding. Note that I add 1 to returns in both the current period and last period. This could avoid both a jump in the optimal manager effort function around $r_{i,t-1} = 0$ and a negative optimal manager effort when $r_{i,t-1} < 0$, both of which will become clear in the derivations in section 4.

3.2. Mutual Fund Investor Decision

I assume there are n homogeneous rational investors in the model. They have the same beliefs, the same preferences, and the same access to all investment opportunities. Therefore, all investors will make the same decision. The value of n affects the model equilibrium through the fund's total net assets and thus the mutual fund family decision, which is treated as a constant in the current version of the model. Hence, both the manager's decision and the decision of each individual CARA investor are independent of the value of n in the current version of the model. Therefore, n does not have any impact on the current version of the model. Mutual fund investors' decision-making process involves two steps. Firstly, the investors observe the most recent fund return at the end of each period and update their beliefs about manager skill in a Bayesian manner. Then, at the end of each period, each investor rationally chooses the amount of capital to allocate to each fund to maximize their expected utility at the end of the next period. Multiplying the optimal amount of capital allocation for each investor by the number of investors, n, gives us the total net asset of the fund predicted by the model. Note that the total amount of capital allocation by investors is not equivalent to the concept of fund flow, which is commonly discussed in the literature. The total amount of capital allocation I am concerned with here refers to the total net asset allocated to each fund, including both the capital investors have already put into the funds in previous periods and the capital generated by the fund as previous returns. The traditional flow measure is the change in the total net asset, net of fund returns. Unlike the implicit assumption in the proportional flow measure, the current period investor capital allocation should be independent of the amount of capital they previously put into the fund. I choose to focus on total net asset as a more straightforward value of interest in this rational model. As argued previously, the total capital allocation variable here is also superior to the proportional flow measure because total capital allocation is robust to Spiegel and Zhang (2013) criticism against proportional flow.

3.2.1. Mutual Fund Investor Belief

At the end of each period, mutual fund investors observe the fund return in this period and use it to update their beliefs about the manager skill using Bayes' rule. I assume the manager skill to be normally distributed. As introduced earlier, the return structure is given as follows:

$$r_{it} = s_i \cdot e_{it} + \epsilon_{it} \tag{7}$$

Because I assume every party knows others' optimal strategy, mutual fund investors know the optimal effort function, e^*_{it}, chosen by mutual fund manager i at time t. As mentioned earlier, the specification of the cost of effort I assume has the benefit of eliminating manager skill, which mutual fund investors do not know for sure, from the manager's optimal effort function. All other variables in the manager's optimal effort function, such as compensation parameters, α_i and β_i, are assumed to be known by investors. Therefore, mutual fund investors treat optimal effort as a constant when updating their beliefs. As a result, the signal, r_{it}, observed by investors is a linear transformation of the normally distributed variable of interest, s_i, with a normally distributed noise, ϵ_{it}. Therefore, the Kalman filter can be applied recursively by investors to update their beliefs about manager skill, and their beliefs will always have normal distributions. The prior investor belief about the manager skill at each time t is the posterior belief at the previous time $t-1$, and is denoted as $s_i \mid r_{i,t-1} \sim N(\mu_{t-1}, \sigma^2_{t-1})$.

3.2.2. Mutual Fund Investor Decision

I assume each mutual fund investor has the choice of investing in a set of mutual funds, P, and a zero-interest riskless asset. After observing returns and updating beliefs at the end of each period t, each mutual fund investor uses the updated beliefs to decide on the amount of capital, X_{it}, he wants to allocate to each mutual fund i at the end of t. The remaining wealth after allocations to mutual funds will be invested in the zero-interest risk-free asset. The allocated capital to each fund i, X_{it}, will generate a gross return, $r_{i,t+1}$, in period $t+1$ for each investor. At the same time, each mutual fund i will charge a fixed percentage annual management fee, a_i, in period $t+1$ on the capital, X_{it}, allocated to it. The annual management fee is assumed to be constant for each fund over time because this percentage is explicitly set out in the fund's prospectus and usually does not change over the fund's lifetime. I assume all funds have no front-end load or back-end load. I assume mutual fund investors have constant absolute risk aversion utility functions. Therefore, at the end of period t, each mutual fund investor chooses capital allocation, X_{it}, for each fund i in the set of mutual funds, P, to maximize his expected utility at the end of time $t+1$:

$$\begin{aligned} &\max_{X_{it}\,\forall i \in P} \; \mathbb{E}_t\left[-e^{-\lambda \cdot W_{t+1}}\right] \\ &\text{s.t.} \quad W_{t+1} = W_t + \sum_{i \in P} \left(X_{it} \cdot (r_{i,t+1} - a_i)\right) \end{aligned} \tag{8}$$

where λ is the investor's risk aversion coefficient, and W_{t+1} is the investor's total wealth at the end of time $t+1$. These parameters are the same for all investors. Note that I do not assume constraints on short-selling and leverage, so I don't constrain the relation between W_t and X_{it}. Returns net of management fees are what matter to investors. The solution to the above expected utility maximization problem is the optimal capital allocation function of each investor, and this function is known to all parties.

3.2.3. Mutual Fund Family Decision

The last party to be considered in the model is the mutual fund family. In the general setup of this model, the mutual fund family chooses the compensation structure, α_{it}

and β_{it}, at the beginning of each period t. α_{it} represents the fixed base payment to the fund manager regardless of the fund performance or fund revenue. β_{it} represents the intensity of the incentive payment, and I assume the incentive payment depends only on fund return rather than on fund size or fund revenue, as discussed earlier. I assume the management fee percentage, a_i, is exogenous and not chosen by the fund family on an annual basis. In the current version of the model, I have simplified the problem by assuming that the fund family i's compensation structure parameters, α_i and β_i, are constant over time and known to all parties. The constant compensation structure can be justified because most variations in fund manager compensation structures exist cross-sectionally across fund advisors rather than over time or at the fund-level (Ma et al., 2019). Although I am currently treating the compensation structure as a constant, it is one of the future directions to extend my model by introducing the fund family optimal decision function into the model. I believe it will enrich the model implications.

The mutual fund family is the company that employs mutual fund managers and offers mutual fund products to investors. The roles of mutual fund companies are more complex and nuanced in reality. The company that hires mutual fund managers to take on fund portfolio management responsibilities is usually called a fund advisor. There might be another fund company that takes on the marketing and advertising responsibilities for the fund. In this model, the mutual fund family is a company that earns a fixed percentage management fee over the total net assets of the fund and pays a fund manager to perform portfolio management duties. Note that I assume at the end of each period, investors decide on the amount of capital allocation for the following period, so the amount of management fee generated in a period is not influenced by the fund performance in the same period. Hence, the compensation structure decision made by the fund family at the beginning of each period will influence both the manager compensation expense in the current period and the fund management fee revenue in the next period. Therefore, a mutual fund family chooses the compensation structure to maximize the difference between the expected fund revenue in the following period and the expected compensation expense in the current period. I assume no discounting over time for fund families. The optimization problem for the fund family at the beginning of time t is given below:

$$\max_{\alpha_{it},\beta_{it}} \mathbb{E}_{t-1}\left[a_i \cdot n \cdot X_{it}^* - \alpha_{it} - \beta_{it} \cdot f(r_{it})\right] \tag{9}$$

where α_{it} and β_{it} are compensation parameters for period t set by the fund family. n is the number of homogeneous investors in the economy. X_{it}^* is the total amount of capital allocated by a single investor to fund i at the end of period t. $a_i \cdot n \cdot X_{it}^*$ is the fund revenue in period $t+1$ because the compensation structure influences the revenue only in the following period through the realized return in the current period t. $\alpha_{it} + \beta_{it} \cdot f(r_{it})$ is the manager compensation paid by the fund family in period t. I introduce the general setup of the fund family decision problem here. For now, I am simplifying the model by treating compensation parameters, α_i and β_i, as constants and known by all parties. In the future version of the model, I plan to solve this maximization problem and add the fund family's optimal decision to the model.

4. Derivations, Simulation, and Further Test

In this section, I will formally solve the model and derive the model implications. To start with, I will consider the manager decision problem.

4.1. Mutual Fund Manager Decision

As discussed earlier, the manager chooses his effort in a certain period at the beginning of that period to maximize the expected end-of-period utility, which is derived from the total wealth minus the current cost of effort. Given the contract specification motivated by empirical analysis in the last section, the maximization problem for manager i at the beginning of time t becomes the following:

$$\begin{aligned}
&\max_{e_{it}} \ \mathbb{E}_{t-1}\left[-e^{-\lambda_i \cdot (I_{it} - \gamma_i \cdot s_i \cdot e_{it}^2)}\right] \\
&\text{s.t.} \quad I_{it} = I_{i,t-1} + \alpha_i + \beta_i \cdot \frac{1 + r_{it}}{1 + r_{i,t-1}}
\end{aligned} \tag{10}$$

where e_{it} is the time t effort chosen by manager i at the beginning of time t. λ_i is the risk aversion coefficient for manager i. γ_i is the cost of effort parameter for manager i. The cost of effort also depends on the manager skill, s_i. I_{it} is the wealth for manager i at the end of time t. α_i is the annual fixed payment to the manager, and β_i is the amount of incentive payment to the manager for each unit of performance improvement over the performance in the last year. Return at time t, r_{it}, is assumed to be $s_i \cdot e_{it} + \epsilon_{it}$. To the fund manager, the only uncertain component in the utility function is the normally distributed noise term ϵ_{it} in the current fund return r_{it}. Therefore, the exponent in the utility function is conditionally normally distributed. Given the CARA utility function and the normally distributed exponent, the maximization problem can be rewritten in the following mean-variance form:

$$
\max_{e_{it}} \mathbb{E}_{t-1}\left[I_{i,t-1} + \alpha_i + \beta_i \cdot \frac{1+r_{it}}{1+r_{i,t-1}} - \gamma_i \cdot s_i \cdot e_{it}^2\right]
- \frac{1}{2} \cdot \lambda_i \cdot \mathrm{Var}_{t-1}\left(I_{i,t-1} + \alpha_i + \beta_i \cdot \frac{1+r_{it}}{1+r_{i,t-1}} - \gamma_i \cdot s_i \cdot e_{it}^2\right)
$$

At the beginning of time t when the fund manager makes the decision, he already observes every term in the above problem, except the current period return. The manager also knows the underlying data generating process for r_{it}. Hence, the manager decision problem can be simplified into:

$$
\begin{aligned}
& \max_{e_{it}} \ I_{i,t-1} + \alpha_i + \beta_i \cdot \frac{1+\mathbb{E}_{t-1}[s_i \cdot e_{it} + \epsilon_{it}]}{1+r_{i,t-1}} - \gamma_i \cdot s_i \cdot e_{it}^2 - \frac{1}{2} \cdot \lambda_i \cdot \mathrm{Var}_{t-1}\left(\beta_i \cdot \frac{1+s_i \cdot e_{it} + \epsilon_{it}}{1+r_{i,t-1}}\right) \\
= & \max_{e_{it}} \ I_{i,t-1} + \alpha_i + \beta_i \cdot \frac{1+s_i \cdot e_{it}}{1+r_{i,t-1}} - \gamma_i \cdot s_i \cdot e_{it}^2 - \frac{1}{2} \cdot \lambda_i \cdot \frac{\beta_i^2}{(1+r_{i,t-1})^2} \mathrm{Var}_{t-1}(s_i \cdot e_{it} + \epsilon_{it}) \\
= & \max_{e_{it}} \ I_{i,t-1} + \alpha_i + \beta_i \cdot \frac{1+s_i \cdot e_{it}}{1+r_{i,t-1}} - \gamma_i \cdot s_i \cdot e_{it}^2 - \frac{1}{2} \cdot \lambda_i \cdot \frac{\beta_i^2}{(1+r_{i,t-1})^2} \cdot \sigma_\epsilon^2
\end{aligned}
$$

where σ_ϵ^2 denotes the constant variance of the noise term in the fund return. Setting the first derivative of the maximization problem to zero, I get the optimal decision function for mutual fund manager i at time t:

$$
e_{it}^* = \frac{\beta_i}{2 \cdot (1+r_{i,t-1}) \cdot \gamma_i} \tag{11}
$$

This solution is essentially the effort level that sets the marginal benefit of effort equal to the marginal cost of effort. The intuition here is that with a lower past return as a benchmark, it's easier for the manager to achieve improvement over past performance, which is the criterion for incentive payment. Therefore, the manager will receive a higher expected benefit for each unit of effort if the past return is lower, resulting in a higher effort choice by the manager. The first thing to notice about this optimal manager effort function is that the optimal effort, e_{it}^*, is a continuous function of past return, $r_{i,t-1}$, as long as the past return, $r_{i,t-1}$, is greater than -1. As mentioned in section 3, if the incentive payment depends on $\frac{r_{it}}{r_{i,t-1}}$, rather than $\frac{1+r_{it}}{1+r_{i,t-1}}$, the optimal manager effort will not be continuous around 0, and managers with negative past returns will have negative optimal effort, which is hard to interpret and contradicts the model assumption. Thus, I choose the specific compensation structure above to avoid such problems. Secondly, the manager's optimal effort does not depend on the manager skill, s_i. This is the benefit of using the specific cost of effort specification $\gamma_i \cdot s_i \cdot e_{it}^2$, because the manager skill variable in this cost of effort function cancels out with the manager skill variable in the return process. With an optimal manager effort function independent of manager skill, the investors' belief-updating process will become straightforward. Thirdly, the optimal manager effort above is decreasing and convex in the past fund return, consistent with the empirical finding in section 2. This simple model specification not only captures the direction but also the convexity of the relationship between the two variables, so I will use this simple contract specification throughout my model. One can also notice that the manager's risk aversion, λ_i, does not appear in the manager's optimal decision function.

I would also like to draw attention to the similarity between the implications of this performance improvement-based incentive structure and the implications of a fund revenue-based incentive structure, which is more prevalent in the mutual fund industry (Ibert et al., 2018). As shown in the optimal manager effort function, the performance improvement-based incentive structure used in this paper generates an optimal manager effort that is negatively correlated with past returns. A fund revenue-based incentive structure would also induce the manager to devote lower effort in the year following a high return. When a fund has a high past return, the fund size in the following year would increase due to both fund return and fund inflow. Because the fund size and fund revenue are independent of the manager effort in the concurrent year, the high fund revenue-based incentive payment in the year following a high return can be regarded as a high fixed base payment, which would reduce the manager's incentive to devote effort in that year due to the income effect of labor supply. Therefore, although I do not adopt a fund revenue-based incentive contract in my model, the model implications generated using my contract specification would also be useful in situations where a fund revenue-based incentive contract is adopted.

4.2. Mutual Fund Investor Decision

As discussed in section 3, the investor decision process consists of two steps: a Bayesian update of beliefs about manager skill and maximization of expected future utility by choosing optimal capital allocation.

4.2.1. Mutual Fund Investor Belief

Mutual fund investors observe fund returns at the end of each period and use Bayes' rule to update their beliefs about fund manager skill. Because of the model specification, the fund return is a linear transformation of the normally distributed manager skill. Thus, Bayesian updating of the investors' belief about manager skill is equivalent to applying the Kalman filter, which allows recursive belief updating and always yields a normally distributed posterior belief.

First, denote the prior investor belief at time t about manager skill as follows:

$$s_i \mid r_{i,t-1} \sim N(\mu_{i,t-1}, \sigma^2_{i,t-1}) \tag{12}$$

The distribution of time t return, r_{it}, conditional on manager skill, s_i, is given by:

$$r_{it} \mid s_i \sim N(s_i \cdot e^*_{it},\, \sigma^2_\epsilon) \tag{13}$$

At the end of period t, investors observe fund returns in period t and apply Bayes' rule to derive the following posterior investor belief about manager skill:

$$f_{S|R}(s_i \mid r_{it}) = \frac{1}{f_R(r_{it}) \cdot 2\pi \cdot \sigma_{i,t-1} \cdot \sigma_\epsilon} \times e^{-\frac{(s_i - \mu_{i,t-1})^2}{2\sigma^2_{i,t-1}} - \frac{(r_{it} - s_i \cdot e^*_{it})^2}{2\sigma^2_\epsilon}} \tag{14}$$

Because the only variable that investors do not know for sure is manager skill, s_i, and manager optimal effort, e^*_{it}, does not depend on the unknown manager skill, the probability density function above represents the posterior investor belief with the following conditional normal distribution:

$$s_i \mid r_{it} \sim N(\mu_{it}, \sigma^2_{it}) \tag{15}$$

where

$$\mu_{it} = \frac{\mu_{i,t-1} \cdot \sigma^2_\epsilon + r_{it} \cdot e^*_{it} \cdot \sigma^2_{i,t-1}}{\sigma^2_\epsilon + \sigma^2_{i,t-1} \cdot (e^*_{it})^2} \tag{16}$$

and

$$\sigma^2_{it} = \frac{\sigma^2_\epsilon \cdot \sigma^2_{i,t-1}}{\sigma^2_\epsilon + \sigma^2_{i,t-1} \cdot (e^*_{it})^2} \tag{17}$$

The mean of the posterior belief is linear to the current observation, r_{it}, as is typical in Kalman filtering. The posterior standard deviation, σ_{it}, is not dependent on return, r_{it}, whereas it is decreasing in optimal manager effort, e^*_{it}. The intuition here is that higher manager effort increases the signal-to-noise ratio in the observed signal, r_{it}, because the noise, ϵ_{it}, in time t return, r_{it}, is independent of manager effort, whereas the meaningful signal, $s_i \cdot e^*_{it}$, is increasing in manager effort. This illustrates the similarity of this model to Franzoni and Schmalz (2017), which also allows the signal-to-noise ratio in past fund returns to vary. Franzoni and Schmalz (2017) do so by allowing variations in the noise component, whereas I do so by having variations in the signal component. Note that the posterior variance, σ^2_{it}, is always less than the prior variance, $\sigma^2_{i,t-1}$, as long as neither the prior variance, $\sigma^2_{i,t-1}$, nor manager effort, e^*_{it}, is zero. This indicates that investors have, on average, more precise beliefs about the skills of fund managers who have been around for longer.

4.2.2. Mutual Fund Investor Capital Allocation

After Bayesian updating, each mutual fund investor will use the posterior belief to make his capital allocation decision. As discussed in section 3, each mutual fund investor chooses from a set of mutual funds, P, and a zero-interest riskless asset in each period. At the end of period t, a mutual fund investor chooses capital allocation, X_{it}, to each fund i in the set of mutual funds, P, to maximize his expected utility at the end of time $t+1$:

$$\max_{X_{it}\,\forall i\in P} \mathbb{E}_t\left[-e^{-\lambda\cdot W_{t+1}}\right] \tag{18}$$
$$\text{s.t.} \quad W_{t+1} = W_t + \sum_{i\in P}(X_{it}\cdot(r_{i,t+1}-a_i))$$

where λ is the investor's risk aversion coefficient, and W_{t+1} is the investor's total wealth at the end of time $t+1$. a_i is the exogenous annual percentage management fee for fund i. $r_{i,t+1}$ is the gross return of fund i at time $t+1$.

To investors at the end of time t, the only uncertain term in the expected utility function is the future return $r_{i,t+1} = s_i \cdot e^*_{i,t+1} + \epsilon_{i,t+1}$. Next period, the manager's optimal effort $e^*_{i,t+1}$ is known to investors in equilibrium because I assume each party knows the strategies of the other parties. Both investors' beliefs about manager skill s_i and future noise $\epsilon_{i,t+1}$ are normally distributed, so future wealth W_{t+1} is also normally distributed. Therefore, we can rewrite the above decision problem in the following mean-variance form:

$$\max_{X_{it}\,\forall i\in P} \mathbb{E}_t\left[W_t + \sum_{i\in P}(X_{it}\cdot(r_{i,t+1}-a_i))\right] - \frac{1}{2}\lambda\cdot \mathrm{Var}_t\left(W_t + \sum_{i\in P}(X_{it}\cdot(r_{i,t+1}-a_i))\right)$$

Substituting the return structure into the maximization problem, the problem becomes:

$$
\begin{aligned}
&\max_{X_{it}\,\forall i\in P}\ \mathbb{E}_t\left[W_t + \sum_{i\in P}(X_{it}\cdot(s_i\cdot e^*_{i,t+1} + \epsilon_{i,t+1} - a_i))\right] \\
&\qquad\qquad\qquad - \frac{1}{2}\lambda\cdot \mathrm{Var}_t\left(W_t + \sum_{i\in P}(X_{it}\cdot(s_i\cdot e^*_{i,t+1} + \epsilon_{i,t+1} - a_i))\right) \\
= &\max_{X_{it}\,\forall i\in P}\ W_t + \sum_{i\in P}(X_{it}\cdot(e^*_{i,t+1}\cdot \mathbb{E}_t[s_i] - a_i)) - \frac{1}{2}\lambda\cdot \mathrm{Var}_t\left(\sum_{i\in P}(X_{it}\cdot s_i\cdot e^*_{i,t+1} + X_{it}\cdot\epsilon_{i,t+1})\right)
\end{aligned}
$$

We assume the noise term follows the distribution $\epsilon_{i,t+1} \overset{iid}{\sim} N(0,\sigma^2_\epsilon)$. Manager skills are also independent across managers. Therefore, future noise, $\epsilon_{i,t+1}$, is independent across funds and independent of manager skill, s_i. Future manager effort, $e^*_{i,t+1}$, is known to investors. Then we can further simplify the maximization problem into the following form:

$$
\begin{aligned}
= &\max_{X_{it}\,\forall i\in P}\ W_t + \sum_{i\in P}(X_{it}\cdot(e^*_{i,t+1}\cdot\mu_{it} - a_i)) - \frac{1}{2}\lambda\cdot\sum_{i\in P}(\mathrm{Var}_t(X_{it}\cdot s_i\cdot e^*_{i,t+1}) + \mathrm{Var}_t(X_{it}\cdot\epsilon_{i,t+1})) \\
= &\max_{X_{it}\,\forall i\in P}\ W_t + \sum_{i\in P}(X_{it}\cdot(e^*_{i,t+1}\cdot\mu_{it} - a_i)) - \frac{1}{2}\lambda\cdot\sum_{i\in P}\left[X^2_{it}\cdot(e^*_{i,t+1})^2\cdot\sigma^2_{it} + X^2_{it}\cdot\sigma^2_\epsilon\right]
\end{aligned}
$$

Setting the first derivatives with respect to the capital allocation, X_{it}, to each fund i to zero, I get each investor's optimal capital allocation function for each fund i at the end of period t:

$$
X^*_{it} = \frac{e^*_{i,t+1}\cdot\mu_{it} - a_i}{\lambda\cdot\left[(e^*_{i,t+1})^2\cdot\sigma^2_{it} + \sigma^2_\epsilon\right]} \tag{19}
$$

Note that this is the optimal capital allocation for each individual mutual fund investor. Since all n investors are homogeneous, the total net asset of fund i at the end of time t, as dictated by the model, is $n \cdot X_{it}^*$. In the optimal investor capital allocation function, both optimal future manager effort, $e_{i,t+1}^*$, and the posterior mean, μ_{it}, are functions of the past fund return, r_{it}, while other variables do not depend on the past fund return. Thus, in my model, the past fund return affects investor optimal capital allocation through two channels: investor belief about manager skill and time-varying manager effort. As in standard Bayesian updating models regarding manager skill, investors apply the Kalman filter after observing the past fund return. Investors adjust their belief about manager skill linearly upwards if the past fund return is high and vice versa. Then, investors will increase the capital allocation if they think the manager has a higher skill in expectation or if they become more certain about the manager skill. The numerator of the investor optimal allocation function, which is the expected future return, is linear with respect to future manager effort, $e_{i,t+1}^*$, whereas the denominator, which is the variance of future return, is a quadratic function of future manager effort. As a result, investor optimal capital allocation is first increasing and then decreasing as a function of future manager effort.

As mentioned earlier, the optimal capital allocation function, X_{it}^*, represents each investor's optimal amount of capital to be allocated to fund i at the end of time t. This value includes both the capital allocated to this fund in previous periods and all net returns already generated by the fund. Thus, X_{it}^* is different from the traditional fund flow measure in the literature. I am going to focus on this capital allocation measure in this paper because it's more straightforward and not subject to common criticisms against the proportional flow measure (Spiegel and Zhang, 2013).

I could have also substituted the expressions for the optimal effort, $e_{i,t+1}^*$, and the posterior beliefs into the individual optimal capital allocation function, X_{it}^*. However, doing so does not bring any convenience because the resulting expression for X_{it}^* will be too complex to read or differentiate. Hence, I am going to numerically simulate the function to visualize the relations between each investor's optimal capital allocation and other variables. Fund total net assets $n \cdot X_{it}^*$ and individual optimal capital allocation X_{it}^* are just linear transformations of each other, so the visualizations of these two variables will have the same shape regardless of the value of n.

4.3. Numerical Simulation

I will now numerically simulate the model to visualize the main implications of the model. The parameters are set to resemble reality and are listed in the following table:

Parameter	Value
$r_{i,t-1}$	0.10
r_{it}	0.12
s_i	0.05
a_i	0.02
α_i	1
β_i	30
γ_i	10
σ_ϵ^2	0.10
$\sigma_{i,t-1}^2$	0.10
$\mu_{i,t-1}$	0.10
λ	0.50
n	1

As shown in the table, I set fund i's return at $t-1$, $r_{i,t-1}$, to 10% and fund i's return at t, r_{it}, to 12%. Note that these returns can be defined in different ways. In the empirical analysis, I define them as Carhart alphas. The true manager skill, s_i, is set to 5%. This means that with each additional unit of manager effort, the expected return of the fund increases by 5%. The annual management fee, a_i, for fund i is equal to 2% of the total net assets. Compensation parameters can have any arbitrary unit. To be realistic, I assume they are in ten thousand dollars. Thus, the fixed payment to the manager, α_i, is \$10,000. Given a β_i equal to 30, if in period t, the manager keeps up with the return in period $t-1$, the manager will earn an additional \$300,000. If the manager doubles this year's return over last year return, the manager will receive \$600,000 as a bonus payment. The cost of effort parameter, γ_i, converts the cost of effort into a monetary value with the same unit as β_i. The variance, σ_ϵ^2, of the noise in the fund return is 10%. The investors' prior belief about manager skill is a normal distribution with both mean, $\mu_{i,t-1}$, and variance, $\sigma_{i,t-1}^2$, equal to 10%. The investors' risk aversion coefficient, λ, is 0.5. I assume there is only one mutual fund investor in the economy. Thus, the fund's total net assets are equal to the investor's optimal capital allocation to the fund. Note that the general shapes of all relations shown below are invariant to the value of n.

Running the simulation with the parameters above, I get the relationship between the total net assets of the fund and the past fund return, as shown in figure 1.

As shown in figure 1, the investor's optimal capital allocation is an increasing and concave function of the past fund return. This coincides with my major empirical finding on the relationship between fund total net assets and past fund abnormal return in table 2. Therefore, my model explains the increasing and concave empirical relation between fund total net assets and past fund abnormal returns as a consequence of rational mutual fund investor behavior. Note that for past returns lower than a certain threshold, the optimal investor capital allocation would be negative. However, since a negative capital allocation to mutual funds is not possible in reality, investors would invest nothing into these very poorly performing funds, and these funds would close down.

The numerical simulation result presented in figure 2 also reproduces the empirical finding of the decreasing and convex relationship between future portfolio activeness and past fund returns.

As shown in figure 2, optimal manager effort is a decreasing and convex function of past fund return, consistent with the discussion in section 4.1. This pattern exactly matches the empirical finding in table 3. Since the expected return in my model is the product of the constant manager skill and the time-varying manager effort, the relation between expected future return and past fund return has a similar pattern, as illustrated in figure 3.

The negative relation between the expected future return and the past return implies that fund returns are not persistent. Note that the expected future return here is also the average future return. This result matches the findings in table 4 in terms of both slope and convexity. It's important to notice that fund performance is not persistent in my model despite the assumption of constant manager skill. Therefore, my model demonstrates that persistent fund performance and the existence of manager skill are not necessarily equivalent, contrary to the mainstream academic argument. My model has a similar spirit to Berk and Green (2004), which also yields constant manager skill and a lack of persistence in fund performance simultaneously. Berk and Green (2004) decouple the manager skill from persistent fund performance by introducing rational investors and diminishing returns to scale, while I do so by having manager effort that is decreasing in past fund returns.

The last implication of the model is illustrated in figure 4. The relationship between the fund total net assets and the future manager's optimal effort, as implied by the model, is shown in figure 4. One can see that if all investors are rational and act optimally, as the model suggests, their capital allocation to the fund should be predicted as a decreasing and convex function of the future manager effort. This empirical pattern is observed in table 2, suggesting that mutual fund investors in the real world are indeed rational in that they correctly expect future portfolio activeness and adjust their investments accordingly.

To summarize, both the direction and convexity of all empirically observed relations in section 2 can be perfectly explained by my rational hidden-action manager effort model.

5. Conclusion

In this paper, I document the novel empirical finding that fund total net assets are an increasing and concave function of past fund abnormal returns. The concave relation between total net assets and past returns presented in this paper is different from the convex or linear flow-performance sensitivity documented in previous literature. I propose investigating the relation between fund past returns and fund total net assets, instead of proportional flow or market share-adjusted flow as in previous literature. Focusing on fund total net assets is straightforward and also avoids Spiegel and Zhang (2013) criticism of the proportional flow measure. This paper serves as a beginning in investigating fund total net assets, and the concave relation presented in this paper needs further research in the future. I also document that total net assets are a decreasing and convex function of future fund portfolio activeness, which is decreasing and convex in past fund abnormal returns. Lack of persistence in mutual fund Carhart 4-factor model abnormal returns is observed in my data.

To explain these empirical findings, I propose a rational principal-agent hidden action model in which mutual fund managers choose their effort levels in response to the incentive structure offered by the fund family and the past fund return, and investors choose the amount of capital allocated to each fund to optimize the mean-variance tradeoff. The Kalman filter is applied by investors to perform a Bayesian update about manager skill. Motivated by my empirical finding on the relation between future portfolio activeness and past fund abnormal returns, I assume a manager payment specification that depends on the magnitude of improvement in current year return over last year return. With this contract specification, the model reproduces all relations documented in the empirical analysis above. This suggests that, in reality, mutual fund managers and mutual fund investors all behave rationally if my model is correct.

This paper contributes to the ongoing academic discussion on mutual fund manager skill, return persistence, and flow-performance sensitivity. I provide a new way to disentangle the concepts of mutual fund manager skill and return persistence by introducing the time-varying manager effort. Similar to Berk and Green (2004), my model generates a lack of persistence in fund returns with constant manager skill assumed, because in my model, managers with higher past returns will devote less effort in the future.

There are a few directions in which to further develop this paper. Firstly, the initial model design involves three parties: mutual fund manager, mutual fund investors, and mutual fund family. In the current version of the model, I have not yet solved the mutual fund family decision, and I am assuming this decision to be constant instead. As the next step, I am going to solve for the optimal compensation structure offered by the fund family and investigate how all three parties interact according to my model. Secondly, I will add a contractible noisy signal about manager effort into the model because, in reality, the portfolio activeness measures can be used as proxies for manager effort. Moreover, I will allow the manager incentive contract in the model to depend on future fund size since Ibert et al. (2018) find that mutual fund revenue is a more significant factor in determining manager compensation than mutual fund return. It will be interesting to see how these three changes made to the model will affect the model implications.

Appendix A. Tables

Table 1: Summary Statistics

This table displays the summary statistics for the data used in this paper. The Carhart model is estimated using monthly observations for each fund in each year. The displayed statistics include the number of observations (N), mean (Mean), standard deviation (σ), 25th percentile (P25), median (Median), and 75th percentile (P75).

	N	Mean	σ	P25	Median	P75
Annual return	58,686	0.12	0.14	0.00	0.12	0.22
Annual turnover ratio	45,929	0.59	0.70	0.26	0.45	0.75
Carhart Alpha	58,441	-0.05	0.06	-0.09	-0.01	0.00
Carhart R-squared	58,441	0.89	0.12	0.83	0.94	0.97
Expense ratio	45,808	0.01	0.01	0.01	0.01	0.01
Fund age	58,827	13.27	9.54	7.00	12.00	17.00
Log TNA	57,536	4.73	1.93	3.47	4.70	6.00
Manager tenure	35,246	8.37	5.65	4.00	7.00	11.00

Table 2: The Determinants of Mutual Fund Total Net Assets

This table presents the fixed-effect panel regressions of mutual fund total net asset on fund past performance and future portfolio activeness. The dependent variable is the total net asset of the fund at the end of the year, in millions of dollars. The independent variables include current-year Carhart 4-factor model abnormal return plus 1 and the future portfolio activeness, proxied by 1 minus the R-squared of Carhart 4-factor model. Quadratic terms of the main independent variables are also included to test for the concavity. The Carhart 4-factor model is estimated using 12 monthly observations for each fund in each year. The control variables are from the same year as the total net asset observations. Control variables include fund annual expense ratio, fund annual turnover ratio, manager tenure, and the age of the fund. All observations are at fund-year level. The sample period is from January 2010 to December 2019. Robust standard errors are computed. Standard errors are reported in parentheses. ***, **, and * denote significance at the 10%, 5%, or 1% level.

	(1)	(2)	(3)
Carhart alpha + 1	18358.304***		19498.582***
	(2936.499)		(3023.673)
(Carhart alpha + 1) squared	-9423.438***		-10107.260***
	(1609.778)		(1649.443)
Future portfolio activeness		-538.761***	-562.498***
		(145.085)	(143.661)
Future portfolio activeness squared		401.882	576.065**
		(278.874)	(285.998)
Expense ratio	-37578.010***	-36952.016***	-36713.674***
	(8540.344)	(8401.786)	(8519.107)
Annual turnover ratio	-37.639***	-38.191***	-36.871***
	(12.976)	(12.777)	(12.925)
Manager tenure	24.112***	24.456***	24.236***
	(6.046)	(6.040)	(6.047)
Fund age	31.218***	31.802***	40.236***
	(7.953)	(7.097)	(8.270)
Constant	-8421.741***	532.311***	-8962.563***
	(1365.288)	(142.154)	(1412.072)
Fund FE	YES	YES	YES
Observations	28350	28510	28350
R^2	0.025	0.025	0.027

Table 3: Mutual Fund Portfolio Activeness and Past Fund Abnormal Return

This table presents the fixed-effect panel regression of fund portfolio activeness on past fund abnormal return. The dependent variable is fund portfolio activeness, proxied by 1 minus R-squared of Carhart 4-factor model. The main independent variable is lagged alpha from Carhart 4-factor model plus 1. The quadratic term of lagged Carhart alpha plus 1 is also included to test for the concavity. The Carhart 4-factor model is estimated using 12 monthly observations for each fund in each year. The control variables are the same as in Amihud and Goyenko (2013), and all control variables are lagged by one year. The sample period is from January 2010 to December 2019. All observations are at fund-year level. Robust standard errors are computed. Standard errors are reported in parentheses. ***, **, and * denote significance at the 10%, 5%, or 1% level.

	(1)	(2)	(3)
Lagged (Carhart alpha + 1)	-1.278***	-0.406***	-1.965***
	(0.010)	(0.015)	(0.523)
Lagged (Carhart alpha + 1) squared			0.842***
			(0.280)
Lagged expense ratio		2.600***	2.576***
		(0.849)	(0.844)
Lagged annual turnover ratio		0.007***	0.007***
		(0.002)	(0.002)
Lagged manager tenure		0.001***	0.001***
		(0.000)	(0.000)
Lagged log TNA		0.006***	0.006***
		(0.002)	(0.002)
Lagged log TNA squared		-0.001***	-0.001***
		(0.000)	(0.000)
Lagged fund age		0.024***	0.024***
		(0.000)	(0.000)
Constant	1.362***	0.163***	0.877***
	(0.009)	(0.021)	(0.244)
Fund FE	YES	YES	YES
Observations	48762	28350	28350
R^2	0.312	0.407	0.407

Table 4: Mutual Fund Return Persistence

This table presents fixed-effect panel regressions that investigate mutual fund return persistence. The dependent variable is the alpha estimated in Carhart 4-factor model. The Carhart 4-factor model is estimated using 12 monthly observations for each fund each year. The main independent variables include lagged Carhart alpha plus 1 and its squared term. The control variables are the same as in Amihud and Goyenko (2013), and all control variables are lagged by one year. All observations are at fund-year level. The sample period is from January 2010 to December 2019. Robust standard errors are computed and reported in parentheses. ***, **, and * denote significance at the 10%, 5%, or 1% level.

	(1)	(2)
Lagged (Carhart alpha + 1)	-14.769***	-12.062***
	(0.201)	(0.198)
Lagged (Carhart alpha + 1) squared	8.585***	6.971***
	(0.109)	(0.108)
Lagged expense ratio		-0.380***
		(0.108)
Lagged annual turnover ratio		-0.001*
		(0.000)
Lagged manager tenure		-0.000***
		(0.000)
Lagged log TNA		-0.000
		(0.001)
Lagged log TNA squared		0.000**
		(0.000)
Lagged fund age		-0.008***
		(0.000)
Constant	6.179***	5.192***
	(0.093)	(0.091)
Fund FE	YES	YES
Observations	48762	28350
R^2	0.907	0.948

Appendix B. Figures

Fig. 1. Total Net Asset-Performance Sensitivity

This figure shows the numerical simulation result of the relation between fund total net asset and fund past return. The parameters are set to the values listed in section 4.3. By plugging in past return values ranging from -20% to 80%, the optimal investor capital allocation is computed according to the model. The optimal investor allocation is plotted against fund past return in this figure.

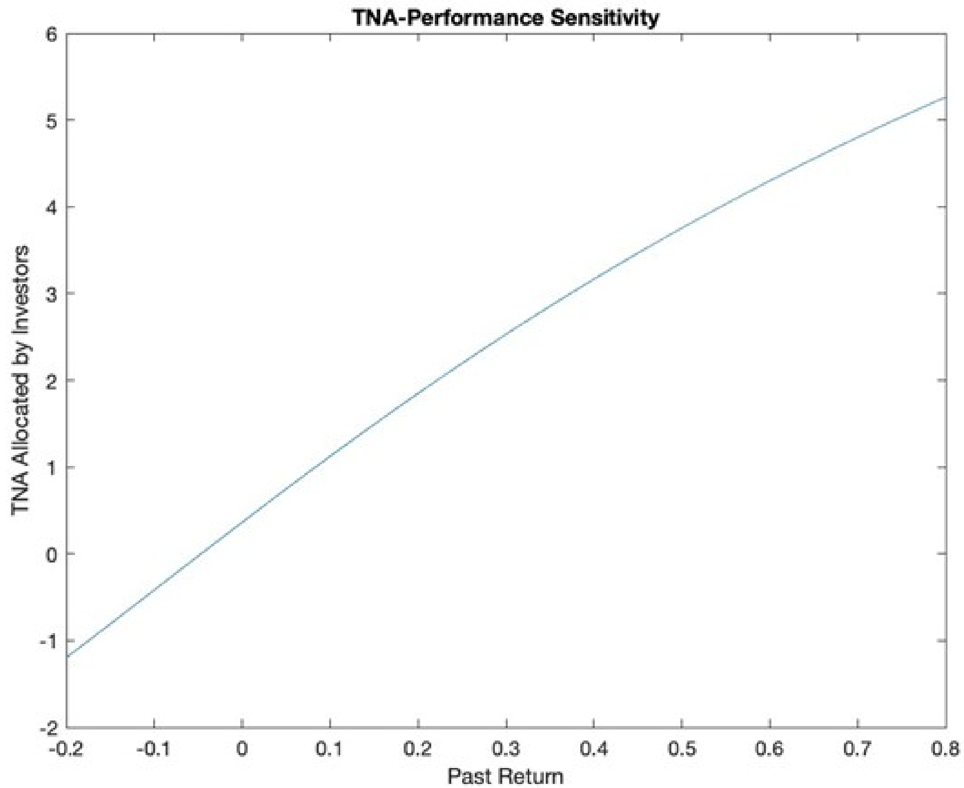


Fig. 2. Optimal Manager Effort and Past Return

This figure shows the numerical simulation result of the relation between current mutual fund manager optimal effort and fund past return. The parameters are set to the values listed in section 4.3. By plugging in past return values ranging from -20% to 80%, the optimal manager effort is computed according to the model. The optimal mutual fund manager effort is plotted against fund past return in this figure.

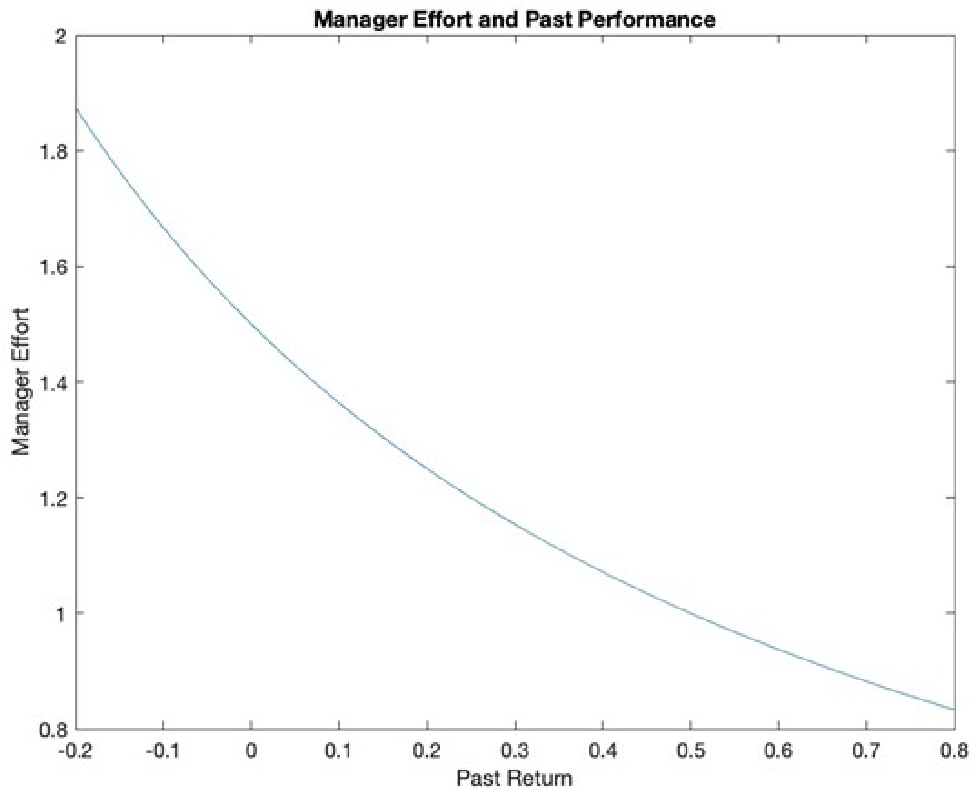


Fig. 3. Mutual Fund Return Persistence

This figure shows the numerical simulation result of the relation between future fund return and past fund return. The parameters are set to the values listed in section 4.3. By plugging in past return values ranging from -20% to 80%, the average future fund return is computed according to the model. The expected future fund return is plotted against fund past return in this figure.

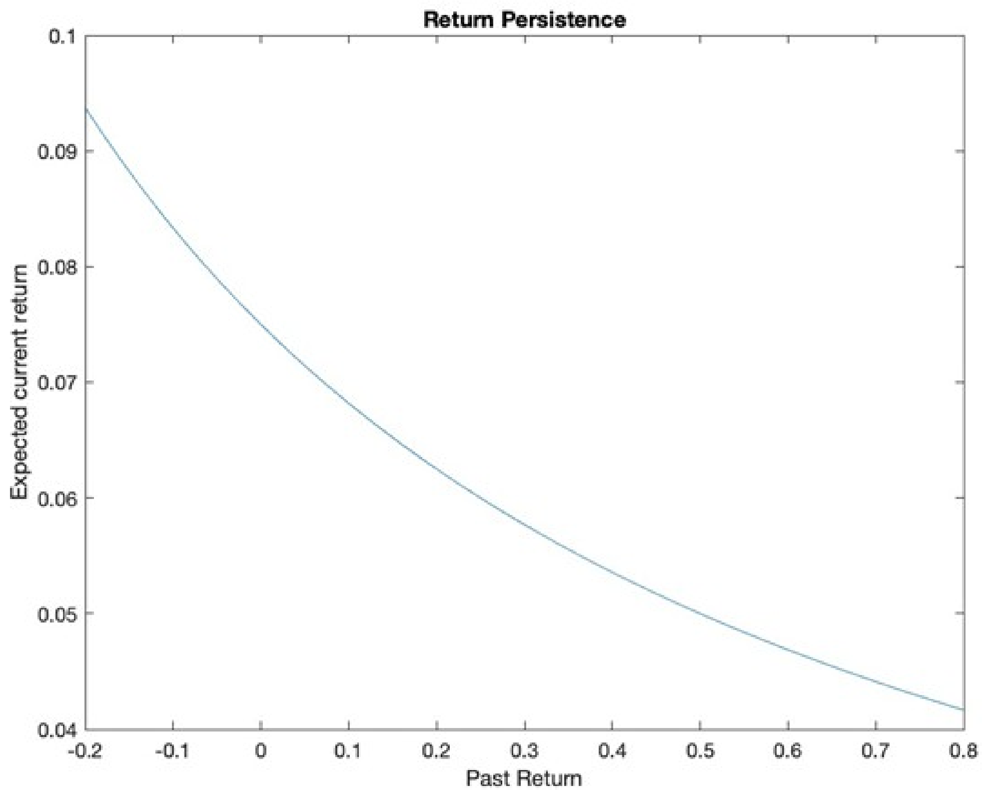
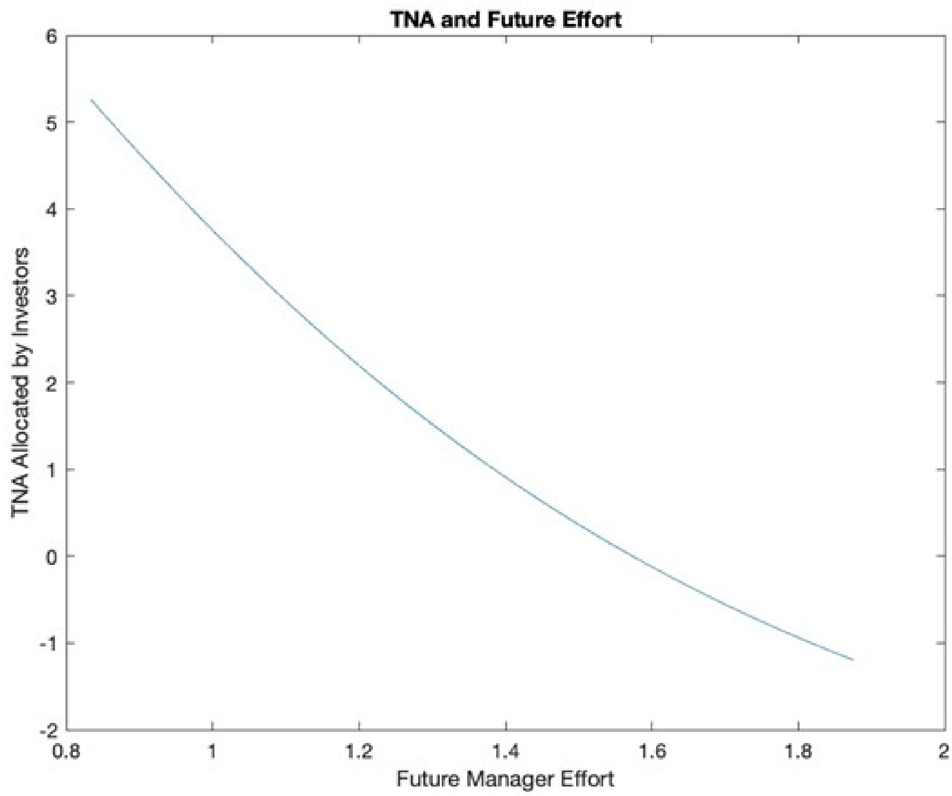


Fig. 4. Fund Total Net Asset and Future Manager Effort
This figure shows the numerical simulation result of the relation between investor optimal capital allocation and future mutual fund manager effort. The parameters are set to the values listed in section 4.3. By plugging in past return values ranging from -20% to 80%, both the optimal investor capital allocation and the optimal future manager effort are computed according to the model. The investor optimal capital allocation is plotted against the future manager optimal effort in this figure.



References

Amihud, Y., Goyenko, R., 2013. Mutual fund's r 2 as predictor of performance. The Review of financial studies 26, 667–694.

Aragon, G. O., Nanda, V., 2012. Tournament behavior in hedge funds: High-water marks, fund liquidation, and managerial stake. The Review of financial studies 25, 937–974.

Berk, J. B., Green, R. C., 2004. Mutual fund flows and performance in rational markets. The Journal of political economy 112, 1269–1295.

Brown, K. C., Harlow, W. V., Starks, L. T., 1996. Of tournaments and temptations: An analysis of managerial incentives in the mutual fund industry. The Journal of finance (New York) 51, 85–110.

Carhart, M. M., 1997. On persistence in mutual fund performance. The Journal of finance (New York) 52, 57–82.

Chevalier, J., Ellison, G., 1997. Risk taking by mutual funds as a response to incentives. The Journal of political economy 105, 1167–1200.

Chevalier, J., Ellison, G., 1999. Are some mutual fund managers better than others? cross-sectional patterns in behavior and performance. The Journal of finance (New York) 54, 875–899.

Cremers, K. J. M., Petajisto, A., 2009. How active is your fund manager? a new measure that predicts performance. The Review of financial studies 22, 3329–3365.

Dangl, T., Wu, Y., Zechner, J., 2008. Market discipline and internal governance in the mutual fund industry. The Review of financial studies 21, 2307–2343.

Ellul, A., Pagano, M., Scognamiglio, A., 2020. Career risk and market discipline in asset management. The Review of financial studies 33, 783–828.

Elton, E. J., Gruber, M. J., Blake, C. R., 1996. Survivorship bias and mutual fund performance. The Review of financial studies 9, 1097–1120.

Elton, E. J., Gruber, M. J., Blake, C. R., 2003. Incentive fees and mutual funds. The Journal of finance (New York) 58, 779–804.

Evans, R., Gómez, J.-P., Ma, L., Tang, Y., 2024. Peer versus pure benchmarks in the compensation of mutual fund managers. Journal of financial and quantitative analysis 59, 3101–3138.

Evans, R. B., 2010. Mutual fund incubation. The Journal of finance (New York) 65, 1581–1611.

Franzoni, F., Schmalz, M. C., 2017. Fund flows and market states. The Review of financial studies 30, 2621–2673.

Hodder, J. E., Jackwerth, J. C., 2007. Incentive contracts and hedge fund management. Journal of financial and quantitative analysis 42, 811–826.

Holmström, B., 1979. Moral hazard and observability. Bell Journal of Economics 10, 74–91.

Holmström, B., Milgrom, P., 1991. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. Journal of law, economics, & organization 7, 24–52.

Ibert, M., Kaniel, R., Van Nieuwerburgh, S., Vestman, R., 2018. Are mutual fund managers paid for investment skill? The Review of financial studies 31, 715–772.

Koijen, R. S., 2014. The cross-section of managerial ability, incentives, and risk preferences. The Journal of finance (New York) 69, 1051–1098.

Lee, J. H., Trzcinka, C., Venkatesan, S., 2019. Do portfolio manager contracts contract portfolio management? The Journal of finance (New York) 74, 2543–2577.

Lynch, A. W., Musto, D. K., 2003. How investors interpret past fund returns. The Journal of finance (New York) 58, 2033–2058.

Ma, L., Tang, Y., Gómez, J.-P., 2019. Portfolio manager compensation in the u.s. mutual fund industry. The Journal of finance (New York) 74, 587–638.

Massa, M., Patgiri, R., 2009. Incentives and mutual fund performance: Higher performance or just higher risk taking? The Review of financial studies 22, 1777–1815.

Ou-Yang, H., 2003. Optimal contracts in a continuous-time delegated portfolio management problem. The Review of financial studies 16, 173–208.

Sirri, E. R., Tufano, P., 1998. Costly search and mutual fund flows. The Journal of finance (New York) 53, 1589–1622.

Spiegel, M., Zhang, H., 2013. Mutual fund risk and market share-adjusted fund flows. Journal of financial economics 108, 506–528.