London School of Economics and Political Science

ESSAYS ON PENSION, INSURANCE AND MUTUAL FUND MARKETS

Jiaxing Tian

Thesis submitted to the Department of Finance of the London School of Economics and Political Science for the degree of Doctor of Philosophy, London, April 2024

Declaration

I certify that the thesis I have presented for examination for the Ph.D. degree of the London School of Economics and Political Science is solely my own work other than where I have clearly indicated that it is the work of others (in which case the extent of any work carried out jointly by me and any other person is clearly identified in it).

I warrant that this authorisation does not, to the best of my belief, infringe the rights of any third party.

I confirm that Chapter 1 is jointly co-authored with Jiahong Shi and I contributed 50% of these works. I also confirm that Chapter 2 is jointly co-authored with Li An, Wei Huang, Dong Lou and Yongxiang Wang and I contributed 20% of this work.

I declare that the thesis consists of 29,137 words.

Acknowledgement

I am deeply indebted to my main advisor, Dong Lou, who guided me throughout my PhD journey with not only academic inspiration and knowledge but also his patient encouragement in my life. He, as a mentor and a friend, sets a brilliant example for me to follow as an academic.

I also express my high appreciation to the other members of my job market committee – Michela Verardo, Christopher Polk, and Cameron Peng. They supported me and provided valuable comments in the first and third chapters of this thesis.

I also learned a remarkable amount from working with Li An, Wei Huang, Jiahong Shi, and Yongxiang Wang, my co-authors in the first and second chapters of this thesis. Their professional experiences and academic views make the works interesting.

This thesis also benefited from conversations and feedback from many other people, notably Charles Cao, Pedro Bordalo, Shiyang Huang, Dirk Jenter, Ralph Koijen, Peter Kondor, Martin Oehmke, and Xuan Tian.

I am grateful to the faculty and staff at the Department of Finance who supported me at various stages, especially Ashwini Agrawal, Cynthia Balloch, Georgy Chabakauri, Mary Comben, Kim Fe Cramer, Vicente Cuñat, Daniel Paravisini, Walker Ray, and Dimitri Vayanos. I acknowledge the financial support of the London School of Economics and the Financial Market Group.

I have been lucky to spend my PhD years with my cohort: Bijan Aghdasi, Guido Maia da Cunha, Yanhuan Tang, Song Xiao, and Sherry Xue. Thanks also go to my colleagues and friends Jingxuan Chen, Juan Chen, Chris Greiner, James Tengyu Guo, Zhongchen Hu, Jiantao Huang, Faqiang Li, Kai Li, Xinchen Ma, Gosia Ryduchowska, Irina Stanciu, Bo Tang, Karamfil Todorov, Su Wang, Hong Xiang, Zheyuan Yang, Zhiqiang Ye, Xiang Yin, and Yue Yuan.

Last, but certainly not least, I thank my fiancee and my parents for their love and support.

Abstract

This thesis contains two essays on investor decisions and the role of financial intermediaries in pension and insurance markets, and one essay on the size effect in the mutual fund market.

In the first chapter, my coauthor and I study how investors respond to scandals related to distinct aspects of environmental, social, and governance in their 401(k) retirement savings. We show that nearby ESG scandals correlate with increased ESG fund additions and flows, possibly through "evoking" existing sustainable preferences among investors. Investors with different characteristics respond heterogeneously to E, S, and G scandals, resulting in an overweighting of funds with higher environmental and social scores.

In the second chapter, my coauthor and I study the impact of sales channels on insurance product adoption. Using novel policy-level life insurance data in China, we exploit a regulatory change that requires bank insurance agents in each quarter to sell more long-term insurance products. Exploiting a discontinuity-in-slope design, we show that bank agents falling below their target qualified ratios in the first two months of a quarter make up for the shortfall in the third month. This shift in the qualified ratio in the last month of the quarter is entirely due to a product-composition change – switching from short-term unqualified life insurance products to long-term qualified annuity products. We further show that this switch is not achieved by changing the relative pricing of products or client compositions.

In the third chapter, I examine the relationship between the magnitude of the negative size effect and fund sector concentration. It finds a strong correlation indicating that funds in more concentrated sectors exhibit more severe diminishing returns to scale compared to those in less concentrated sectors. The paper proposes a potential explanation: in highly concentrated sectors, fund flows are less sensitive to past returns. However, in such sectors, marketing expenses appear to positively influence flow sensitivity to good performance, while showing a neutral effect in response to poor performance. Large funds in concentrated sectors may invest more in marketing efforts, but this does not necessarily translate to better future performance.

Contents

List of Tables

List of Figures

Chapter 1

Unpackaging ESG: Evidence from 401(k) Investment

Jiahong Shi and Jiaxing Tian[1]

We study how investors respond to scandals related to three distinct aspects of ESG–E(nvironmental), S(ocial) and G(overnance)–in their retirement savings. Using data on 401(k) investments, we show that nearby ESG scandals correlate with increased ESG fund additions and flows, possibly through "evoking" their existing sustainable preferences. Investors with different characteristics respond heterogeneously to E, S and G scandals. In magnitude, old investors are twice as likely as young investors to add ESG funds to their portfolios after the shock of social scandals. In specific scandals, low-income investors care about human rights, while only young and rich investors care about environmental issues. Investors also have clear leanings on ESG funds, resulting in an overweighting of funds with higher environmental and social scores and a lack of attention to governance elements. Overall, the results suggest the need to incorporate distinct E, S, and G concerns into heterogeneous preference models.

[1]This is the Jiaxing Tian's job market paper. We thank Dong Lou, Christopher Polk, Michela Verardo, Cameron Peng, Dirk Jenter, Pedro Bordalo, Shiyang Huang, Dimitri Vayanos and seminar participants at CUHK-Shenzhen, NEOMA Business School, PBCSF Tsinghua University, Peking University HSBC Business School, and Renmin University for helpful comments. We thank Financial Market Group for the financial support. All remaining errors are my own.

1.1 Introduction

Sustainable concerns have been taken into consideration in financial investment. Investments in ESG (Environmental, Social, and Governance) funds have rapidly grown over the past decades, exceeding $236 billion in the U.S. by the end of 2020. Despite the large size of ESG product market, the ESG product market is "homogeneous". On the one hand, approximately 50% of ESG funds are named with overbroad terms like "Social Responsible Fund" or "Green Investment Fund". More than 80% of ESG funds use three main ESG indices, namely MSCI ESG Index, S&P500 ESG Index, and FTSE Russell ESG index, as benchmarks. On the other hand, the number of specifically targeted funds, such as consumer rights funds and gender equality funds, which do not follow these indices, is limited, with around 50 funds of $35 billion in AUM (Assets Under Management) in the U.S. by the end of 2020.

People invest in ESG funds, particularly in their pension investments, but their preferences may exhibit heterogeneity. The heterogeneity could come from the fact that investors care about different economic implications of environmental, social and governance concerns. Firm environmental and social issues are relevant to externalities due to market failure, while governance issues are internal agency problems. Investors may value these concerns separately, while asset management company usually puts these three aspects into one product. These concerns are expanding, but many newborn issues, including data privacy and greenwashing, are less likely to be covered in the ESG product design than old issues like pollution. Investors may have difficulties finding products that align perfectly with their preferences in E, S, and G. Consequently, conflicts may arise between the "homogeneous" product market and the heterogeneous preferences.

To understand heterogeneous preferences, the main difficulty is to find revealed preferences, given investors are holding homogeneous products. In this paper, we study how investors make pension investment decisions in response to E, S, and G scandals. Investor decisions are cross-matched with demographic features, including political leaning, age, and wealth. We use ESG scandals that happened in the previous year and near the plan location as shocks to investor portfolio changes.

These scandals could be categorized into specific issues representing different ESG concerns, including human rights, biodiversity, technology, etc. We focus more on the comparable magnitude of heterogeneous responses, which could serve as the first step towards calibrating heterogeneous investor models. This question also helps guide ESG indices and fund setting-ups.

Although there are multiple channels to make investors change their portfolios after ESG shocks, studying the portfolio changes after ESG shocks could help us understand their sustainable preferences. From asset allocation perspectives, shocks may change the expectation for future cash flows, the inference of systematic risks, the taste of risk tolerance, and non-pecuniary or sustainable utilities. It is most likely that scandals "evoke" latent non-pecuniary concerns. The scandals do not change their preferences, either in the pecuniary part or non-pecuniary part, but make them explicit. We will discuss the mechanism later.

To broadly cover investor heterogeneity, we utilize 401(k) plan investments in the U.S. and assume the nearby ESG scandals from last year would affect this year's 401(k) investment. The defined contribution 401(k) plans cover approximately one-sixth of the U.S. population. We assume that scandals involving a firm primarily affect the area near its headquarters, and plan participants predominantly reside around their sponsor's address.[2] Therefore, we aggregate ESG scandals based on the first two digits of their zip codes to match each plan's location[3]. We rely on publicly accessible data to obtain political leanings from county-level presidential voting records, estimate investor ages from target-date fund investments in each plan, and assess wealth based on average investor contributions. Unlike survey studies, which rely on a limited number of investors or use demographic features in a particular area to represent all investors in that area (Baker, Egan and Sarkar, 2022), our age and wealth estimations capture cross-sectional differences among each plan (firm).

With a one standard deviation increase in the nearby scandal occurrence, the probability of adding an ESG fund would increase by 0.15% to 1%, which is 10–

[2] The robustness of this assumption is further explored in our Appendix.

[3] The scandals involve public firms, and our 401(k) investments include both private and public firms.

70% of the probability standard deviation, and the inflow to ESG funds would increase by 5 to 10 basis points of the total plan assets last year, which is 20–40% of the flow standard deviation. Investors mainly react to unexpected, new, and highly spread scandals, which provide evidence that the scandals evoke their non-pecuniary preferences. Scandals with the highest severity do not contribute more than average to portfolio changes, which suggests scandals are less likely to change their expectations in cash flows or risk evaluation. Psychology studies (e.g., Tarrant, North and Hargreaves, 2001) suggest that social categorizations could trigger latent preferences. Shocks could also be used to distinguish between latent homophily and heterophily (e.g., Ma, Krishnan and Montgomery, 2015).

In menu-change cases, we find clear heterogeneity in all dimensions of characteristics and categories of scandals. We differentiate between unconditional and conditional ESG fund additions, focusing on all plans or just those with menu changes. In both cases, ESG scandals increase the likelihood of ESG fund inclusion within 401(k) plans, but the propensity varies across political leanings, age groups, and wealth levels. Old investors are twice as likely as young people to add ESG funds to their portfolios when facing social scandals. Only young and rich investors care about environmental issues. Rich investors, who also respond to technology and health scandals, are more robust than poor investors, who care about human rights and labor scandals. Democratic investors add ESG funds when facing all kinds of social scandals, especially data privacy, while republican investors marginally react to economic and political scandals to change their portfolios. Given the internal impact of governance issues, investors do not react to governance scandals.

Regarding the effect of ESG scandals on existing fund flows, the inflow is driven by Republican-leaning, younger, and wealthier individuals in response to environmental and social scandals. The results are consistent with the menu-change case, and issues with indirect impacts, like biodiversity and consumer rights, are valued by investors. Although Republican investors rather than Democratic ones induce fund inflow, this could be explained by the heresy status of plans in Republican areas but with ESG funds on their menu, which evidences the higher willingness to pay for ESG. To address potential biases from sample selection, we use the Heckman (1979), sample-selection model, revealing that scandals impact fund inflows only indirectly through the presence of ESG funds in the investment menu.

We show that investors significantly lean towards specific E/S/G elements in the funds even when facing "homogeneous" products. Focusing on those ESG funds in the 401(k) menu, we find a consistent pattern of overweighting social scores and environmental scores by 20–50% in ESG fund-addition decisions after corresponding scandals. When considering ESG fund flows, we find that ESG funds with higher social scores experience a 50% higher inflow when facing incidents compared to other ESG funds. These results provide evidence that investors tend to decompose indices when allocating their assets. While this decomposition helps mitigate the inconsistency between heterogeneous preferences and the composite product, the investor's non-pecuniary utility could not be fulfilled.

Primarily, this paper contributes to understanding heterogeneous preferences in ESG investments. Existing literature often focuses on heterogeneity in either individual elements of ESG or composite ESG concepts. Studies like Hartzmark and Sussman (2019) and Pástor, Stambaugh and Taylor (2021)[4] treat ESG investment decisions and motivations as integrated concepts, without assessing the specific roles of each element within ESG. On the other hand, Bolton and Kacperczyk (2021), Bernstein, Gustafson and Lewis (2019), and Chava (2014) have emphasized environmental shocks in sustainable investment. However, several studies, including Humphrey et al. (2021) and Heeb et al. (2023),[5] have provided evidence that environmental, social, and governance aspects inherently differ in terms of willingness to pay and their social impacts. We demonstrate the magnitude of differences among these three aspects in both investment motivations and decisions, thus highlighting a demand for specific ESG products.[6]

Our work is also dedicated to studying the role of negative ESG shocks in real investment decisions.[7] Ethical concerns and adverse experiences have been exten-

[4]Also see Bauer, Ruof and Smeets (2021).

[5]Also see Riedl and Smeets (2017) and Barber, Morse and Yasuda (2021).

[6]Other works that have supported the impact of ESG demand on fund manager decisions include Alok, Kumar and Wermers (2020), Li et al. (2023), Heinkel, Kraus and Zechner (2001), and Goldstein et al. (2022). We also contribute to evaluating composite ESG products and indices partially, linking to works by Cohen, Gurun and Nguyen (2020), Wu, Zhang and Xie (2020), and Christensen, Serafeim and Sikochi (2022).

[7]Some works propose a wide range of drivers for ESG investments, including the concept of "voice" for addressing externalities. See the works including Broccardo, Hart and Zingales (2022), Broccardo, Hart and Zingales (2022), Berk and Van Binsbergen (2021), and Allcott, Gentzkow

sively studied as specific triggers for ESG investments.[8] Notably, real environmental experiences have been identified as common triggers for economic transformations. As demonstrated by Engle et al. (2020), investors engage in hedging strategies in response to climate change-related news.[9] Our study shifts to board-level ESG-related incidents. Rather than showing the psychological reasons or the pecuniary concerns, we focus on the heterogeneous roles of the shocks. By categorizing aggregated scandals, we compare specific issues and quantify differences across scandal categories in motivating ESG investments.

In addition, this paper contributes to estimating the general investor heterogeneity. Cohen and Einav (2007) estimate the heterogeneous risk preferences, and Xiouros and Zapatero (2010) show that the heterogeneous risk aversion could link to some formation triggers. Based on this, other papers like Chapman and Polkovnichenko (2009) study the outcomes of such investor heterogeneity. This is parallel to our study, and we move to the non-pecuniary part of the heterogeneity.

Lastly, focusing on 401(k) plans, our work contributes to academic research on retirement savings, specifically in portfolio choice. Previous studies often either treat menu design as given or examine the bargaining outcomes between plan sponsors and asset managers.[10] Sialm, Starks and Zhang (2015*b*) and Sialm, Starks and Zhang (2015*a*) discuss the pecuniary reasons that affect 401(k) menu choices and fund flows. Our study extends this analysis to explore the non-pecuniary factors influencing menu changes and asset reallocation.[11] While ESG funds in our sample constitute a small proportion of the market share, we acknowledge their relatively minor market impact and instead offer additional fund configuration recommendations from the demand side.[12]

This paper is structured as follows: In Section 2, we outline the data and the

---

and Song (2022).

[8]e.g. Renneboog, Ter Horst and Zhang (2011), Bernstein, Gustafson and Lewis (2019) and Döttling and Kim (2022)

[9]Other works investigating the effect of real-world experiences on investment decisions include Murfin and Spiegel (2020), Baldauf, Garlappi and Yannelis (2020), Liao et al. (2021), Bradbury, Hens and Zeisberger (2015), and Gompers et al. (2005).

[10]Huberman and Jiang (2006), Davis and Kim (2007), Cohen and Schmidt (2009), Pool, Sialm and Stefanescu (2016), Tang et al. (2010), Pool, Sialm and Stefanescu (2022).

[11]Agnew, Balduzzi and Sunden (2003) and Choi, Laibson and Metrick (2002) demonstrate trading behaviors in 401(k) plans driven by pecuniary concerns.

[12]Other studies related to retirement savings, albeit less directly connected to our work, include Benartzi and Thaler (2001), Choi (2015), and Sialm, Starks and Zhang (2018).

construction of key variables. Section 3 delves into scandal-induced 401(k) menu changes relevant to ESG funds, with a specific emphasis on investor heterogeneity. Section 4 breaks down ESG scandals and investigates their varying impacts across distinct investor groups. Next, in Section 5, we extend our analysis to existing ESG fund flows. Building on these results, Section 6 presents a quantitative analysis of heterogeneous ESG preferences, illustrating how investors may decompose the indices to make investment decisions. Finally, Section 7 encapsulates our conclusions.

1.2 Data and Variable Construction

1.2.1 401(k) Investment

Our primary data sources consist of the Department of Labor (DOL) and BrightScope Beacon. Under regulatory requirements, employers (plan sponsors) are obligated to annually submit Form 5500 to the Department of Labor if their defined contribution plans have at least 100 employee participants at the start of a plan year. Our analysis specifically targets 401(k) plans. Form 5500 captures plan-level details, including location, participant count, service providers, and financial data (Schedule H). Investment details, specifically the assets within individual investment vehicles, are found in Form 5500's appendix. BrightScope Beacon, covering more than 90% of 401(k) defined contribution plans, consolidates plan-level investment data from Schedule H and Form 5500's appendix. The cumulative investment across each option is aggregated across all plan participants. It is important to note that because individual participant holdings remain unobservable, we focus on interplan variations, treating each plan as a representative investor.

Other plan-specific attributes are directly extracted from Form 5500. The dataset spans over 70,000 distinct 401(k) plans across the 2012–2019 period. These plans include both private and public firms, no matter large or small. If a plan invests in the company's stock, then it also must report to the SEC with Form 11-K. Plans filing 11-Ks are just a small subsample[13] of our data. Our analysis selectively

[13] Around 3200 plans are from public firms with 11-K reports.

concentrates on plans with an average asset value exceeding $10 million from 2012 to 2019 and possessing a minimum history of five years post their initial ERISA report. This selection criteria results in a final sample of 29,000 plans, yielding an average of approximately 14,000 observations per annum.

We integrate investment menu data from BrightScope with CRSP, the Morningstar mutual fund database, and supplementary information from DOL Form 5500. On average, each investment menu includes approximately 28 options, with over 25 of them being mutual funds or direct filing entities (DFEs). DFEs refer to mutual fund-like trusts or insurance company separate accounts that directly file and report to the DOL rather than the SEC. CRSP provides monthly returns and total net assets (TNA) data for all open-end funds, while the remaining fund characteristics are from the Morningstar database. Although not covered by CRSP, a portion of DFEs voluntarily disclose information to Morningstar, allowing us to map and obtain returns and holdings for them from Morningstar. For the remaining DFE options without detailed public disclosure, we associate them with mutual funds bearing the same name within the same management family, where possible. We base this on the assumption of similar asset holdings, which lead to comparable gross returns.

Identification of ESG Funds

To identify ESG investments within each 401(k) investment menu, we employ ESG label identification on fund names.[14] If an investment vehicle's name includes certain ESG-related terms like *ESG*, *Green Energy*, or *Human Rights*,[15] it is categorized as an ESG vehicle. In the subsequent discussion, we refer to these ESG investment vehicles as "ESG funds" for the sake of simplicity, despite not all of them being mutual funds. To avoid misinterpretation, we treat funds without ESG-specific names but with a Morningstar ESG rating above 4 as ESG funds as well. Conversely, funds with ESG-related names but a Morningstar ESG rating below 3 are excluded from ESG funds.[16] Because Morningstar ESG ratings were

[14] Following Hartzmark and Sussman (2019), ESG labels matter more for retail investors.

[15] For a comprehensive list of terms, please check the Appendix.

[16] This approach is further validated by the fact that nearly 90% of ESG-named funds exhibit ESG scores for their value-weighted holdings surpassing the median of their respective Morn-

introduced in 2016, we match fund observations before this date with funds of the same name and category after 2016 to infer their pre-2016 ESG rating.

Based on the investment menu of each plan in BrightScope Beacon, the key variable ESG_Funds is the number of ESG investment vehicles in the fund over the total number of options. A menu change could include adding, deleting, or replacing a fund.[17] By comparing two years' observations of a menu, we construct the variable ESG_Add as the number of ESG funds added to the plan in a year over the number of investment options at the end of the previous year. Figure 1.1 illustrates the growth of 401(k) plans offering ESG options, increasing from 1,500 to 2,200 over a decade, with over 100 plans adding ESG funds every year. In total, 401(k) plans invest over $1 billion in ESG funds, as shown in Figure 1.2b. The count of ESG funds available in 401(k) plans has risen from under 100 to nearly 200, as depicted in Figure 1.2a. These 401(k) assets represent approximately 10% of the total net assets of these funds. Table 1.1 shows that, in general, 12% of the investment options are ESG funds, and a plan would add 0.013 ESG funds each year. We do not treat funds closing or merging as either additions or deletions.

To derive distinct E/S/G (Environmental, Social, and Governance) scores for each ESG fund, we integrate Morningstar's quarterly data on mutual fund and commingled trust holdings with 401(k) ESG investment choices. We leverage MSCI ESG ratings, which encompass a comprehensive array of public companies, and merge them with fund-level holdings. We match the E/S/G scores of company stocks from the prior fiscal year with the current year's quarterly holdings. Funds and trusts in our analysis should allocate over 70% of their assets in equities, with 50% of these equities having available MSCI ESG ratings. Fund-level E/S/G scores are computed by aggregating the value-weighted specific scores across the preceding 12 quarters. To mitigate concerns about tracking errors based on diverse fund benchmarks and categories, we also compute normalized E/S/G scores for each fund by subtracting the corresponding average E/S/G scores of each category. For accurate computation, we require a minimum of ten reports for each fund's calculation and a minimum of five funds per category for each quarter. At the plan-year level,

ingstar categories.

[17] Replacing a fund is usually determined by the sponsor to have a fund replaced by a same-category fund due to share class reasons.

E/S/G investment in ESG funds is determined by the value-weighted summation of all investment options at the end of the final quarter of each sample year.

1.2.2 ESG Scandals

To capture the impacts of ESG negative shocks to 401(k) investment, we utilize RepRisk's dataset of 51,000 scandals involving 29,000 firms in the sample period from 2011 to 2019. RepRisk classifies these occurrences into four main issues: environmental, social, governance, and cross-cutting, encompassing a total of 28 sub-issues and 73 specific topic tags. We manually aggregate these 73 tags into 13 issues[18] to study the heterogeneity of scandals in a concise measure.

To maintain objectivity, our analysis focuses on "sharp" incidents.[19] We merge RepRisk data with Compustat and Datastream, resulting in a dataset comprising 5,000 companies and 35,000 scandals. Approximately 85% of the companies in our sample are publicly traded.

Since we only have the exact addresses of each scandal firm's headquarters, we assume that scandals primarily impact the headquarters' location. This assumption does not imply that scandals necessarily occur in that area, but rather that their impacts on ESG are primarily taken by residents near headquarters. Scandals typically have real social impacts, but investing in ESG due to scandal shocks could be driven by attention effects as well. We am not distinguishing the differences between these, and our analysis potentially supports both, which leads to the fact that scandals have the most significant impact around the headquarters of involved firms.[20]

[18]Animal & Biodiversity; Cyber & Privacy Concerns; Economic & Consumer Rights; Environmental Issues; Health & Safety; Labor & Employment; Legal & Compliance; Political & Governance; Resource & Infrastructure; Social & Human Rights; Substance & Social Ethics; Technology & Surveillance; Weapons & Security. For a comprehensive mapping between the issues and the 73 tags, refer to the Appendix.

[19]RepRisk defines unsharp incidents as instances where the entity is mentioned but the criticism is not precisely defined due to the nature of the report.

[20]In our appendix, we present results showing that scandals causing more increased attention are more linked to ESG investment decisions.

1.2.3 Demographic Characteristics: Political Leaning, Age, and Wealth

Our study relies on plan-level data from DOL Form 5500 and the BrightScope database. While these sources offer a comprehensive view of long-term investment, particularly in ESG, they inherently lack the granularity of individual employee characteristics. We treat each plan as a representative investor and employ proxy variables to capture cross-plan differences in demographics. This approach might appear unconventional; however, it helps mitigate idiosyncratic employee-level outliers as well. We capture the plan-level variations in average political leaning, age, and wealth as demographic heterogeneity among investors.

Political Leaning We adopt presidential voting outcomes as a proxy for political inclination. Using county-level voting outcomes from 2012, 2016, and 2020 and weighting the results by voter numbers, we aggregate voting shares to the level of the first two digits of zip codes. In cases where counties consist of multiple such zip code areas, the voting outcomes are used to cover all those respective areas. The annual political leaning is computed through linear interpolation between each of the two closest voting results. Using the first two digits of zip codes is preferable, as plan participants might not reside in the plan sponsor's five-digit zip code address; thus, we assume they are within the same state and share the same two-digit area.[21] The interpolation helps mitigate bias induced by swing areas, ensuring a smooth transition over time. Plans located in the same first two-digit area possess identical variables R_Vote and D_Vote, representing the Republican and Democratic party voting percentages, respectively. A plan is categorized as Democratic-leaning if, in a given year, $D\_Vote > R\_Vote$.

Age To account for investor age variation among plans, we leverage the investment in target-date funds (TDFs) in 401(k) plans. Since 2006, TDFs have become the default investment option in 401(k) plans. These funds transition investment gradually from equity to fixed-income securities as the target retirement year approaches, known as the TDF year. Assuming a retirement age of 65, investors

[21] This is also based on the assumption that employees reside within the two-digit zip code area of the plan's (firm's) location. Given that small and private companies dominate in our sample, this assumption is acceptable.

allocate their investments to TDFs with the TDF year closest to their projected retirement year. For instance, an individual born in 1988 would invest 60% of their money in TDF 2055 and 40% in TDF 2050. However, this approach introduces a bias in estimating actual ages because of wealth accumulation. Despite this bias and overestimation of overall age, the TDF-based age estimate serves as a consistent proxy for cross-sectional age differences across plans. We calculate the TDF-derived age as $TDF\_Age = t - TDF\_Year + 65$, where TDF_Year is the value-weighted average targeting year of all TDFs in the given plan. As indicated in Table 1.1, the median TDF age appears to be 49. In our regression analysis, we incorporate the logarithm of TDF_Age as a plan-specific variable denoted by Log_Age. Another measure to mitigate the impact of wealth accumulation involves the fund flow into TDFs. We classify TDFs into two categories based on their target years—before and after 2040.[22] The flow for each category is computed as follows:

$$
\begin{aligned}
Flow_{TDF \leq 2040,i,t} &= \sum_{k \leq 2040} \frac{V_{k,i,t} - V_{k,i,t-1} R_{k,t}}{V_{k,i,t-1}} \\
Flow_{TDF > 2040,i,t} &= \sum_{k > 2040} \frac{V_{k,i,t} - V_{k,i,t-1} R_{k,t}}{V_{k,i,t-1}}
\end{aligned}
$$

where $V_{k,i,t}$ is the monetary amount invested in the TDF fund with a target year k in plan i during year t, and $R_{k,t}$ represents the return of that particular TDF. The ratio of the two flows, $Flow_{TDF \leq 2040}/Flow_{TDF > 2040}$, highlights the difference in investment growth between older and younger employees. A higher ratio indicates greater flow by older employees. We assume that individuals of the same age exhibit similar investment tendencies toward TDFs and that the investment gaps between old and young employees are similar across different plans. This variable measures the number of older employees relative to younger employees and captures cross-sectional differences. Such a ratio has a mean of 1.22 and a median of 1, according to Table 1.1, but the standard deviation is large, reflecting the heterogeneity of company age structures.

Wealth By employing the average employee contribution, *Deferral*, and the average account balance, $Account\_Balance$, we measure the variations in participant

[22] The year 2040 is based on the median of TDF_Age and is adjusted for the overestimation.

wealth across plans. Assuming a strong correlation between salary and employee wealth, and that employees maximize their vesting amounts within their retirement plans, the deferral amount becomes a stable proportion of their salaries and wealth. Although employer matching varies across plans, this variation is generally moderate.[23] However, it could underestimate the wealth of younger participants and overestimate that of older ones. This is because younger individuals may choose not to maximize their vesting amounts due to intratemporal financial constraints. Consequently, a plan with a higher proportion of young participants might yield an underestimated average wealth. Thus, the age distribution within a plan may correlate with the deferral proxy for wealth. On the other hand, the average account balance is a more straightforward measure but is affected by wealth accumulation and different risk preferences. Plans with higher equity investments in previous years exhibit higher account balances, assuming all other wealth aspects are constant. The plan-level variable *Deferral* is the total deferral of the plan, and *Account_Balance* is the total net asset, both divided by the number of active participants.

1.3 Heterogeneity in Menu Changes: ESG Fund Addition

In this section, we explore 401(k) ESG investment heterogeneity in terms of menu changes. We focus on the aggregated scandals as triggers and the corresponding impacts across investors with different demographic characteristics.

1.3.1 Static Differences in ESG Fund Preference

We first show the unconditional variation in ESG fund inclusion within investment menus. We aggregate the average number of ESG funds across plans grouped by political leaning, age, wealth, and ESG scores of nearby public firms. Table 1.2 presents the group means and differences with t-statistics. This illustrates the

[23] Around 3% for 60% of all plans, and the variation could be considered negligible.

collective equilibrium shaped by influencing factors and also tests the correlations between these characteristics and ESG investments. Specifically, our findings indicate a higher probability of ESG funds being included in the menu within plans characterized by Democratic leanings, younger participants, and higher wealth levels.

The first row of Table 1.2 shows the ESG fund inclusion with respect to the ESG scores of nearby public firms. The variable *Geo_ESG* is defined by aggregating the ESG scores of public firms within the first two digits of their zip codes, weighted by market value, and then matching that with the plan's zip codes. We partition the geographic ESG scores into two distinct groups, annually rebalancing each. Although the 11-K dataset offers insight into ESG investments among retirement plan participants within public firms, it represents only a subset of 401(k) plans, specifically those involving public firms and participants investing in company stocks. If the difference in ESG investments could be fully attributed to this index, our emphasis on the entire 401(k) spectrum would be unwarranted. However, the results show that plans in low-geographic-ESG areas are more likely to have ESG funds in the menu, but the difference is marginal.

The following rows illustrate differences across three key characteristics. Plans in Democratic areas have 4 percentage points more ESG funds in their menus compared to Republican areas. This difference is the most significant among all the features. Age, as measured by TDF flows mentioned above, shows a significant disparity of 2.2 percentage points, with young investors having more ESG funds in their menus. Menus with more-affluent individuals have 2.18 percentage points more ESG funds than the less affluent group.[24]

1.3.2 Menu Changes in Response to ESG Scandals

To study the latent preferences in ESG, we employ ESG scandals as shocks to trigger ESG investments. These scandals are aggregated based on the first two digits of zip codes under the assumption that they have the most significant im-

[24] To control for additional plan-level factors in detecting the relationship between demographic characteristics and ESG investment, we conduct a regression in our Appendix. It shows that political leaning and age are the most robust features.

pacts within the geographic area around the associated firm's headquarters. we do not argue that ESG scandals predominantly occur near the firm's headquarters since we cannot verify the actual locations of the scandals. Instead, we consider ESG scandals as both real-life and salience shocks. Investors residing near the scandal firm's headquarters are more likely to notice the scandal's real impact and also be affected by it, either directly or indirectly. For example, a social scandal related to bribery may have a broader social impact that is challenging to assess in physical locations, but local media would provide more coverage for such a scandal compared to remote media. This implies that local residents are more affected by the scandals. The validity of this assumption is demonstrated in the Appendix. We show that unexpected and time-detrending scandals have more robust and significant results. Indeed, investors with higher exposure to the social impact of scandals react more strongly, as supported by the analysis of firms without any branches.

By summing and normalizing the count of scandals to a scale of one hundred, we use the count of scandals occurring in the preceding year ($ESG\_Scandals_{i,t-1}$) as a key independent variable in our regression analyses. Plans in the same area with the same first two digits of zip code face the level of scandal occurrence. With the same independent variable, we capture the heterogeneous preference for plans with various features. Table 1.3 presents regression coefficients illustrating the impact of ESG scandals on menu changes. The initial column of the regression model is formulated as follows:

$$ESG\_Fund_{i,t} = \beta ESG\_Scandal_{i,t-1} + \mathbf{\Gamma} \boldsymbol{Controls_{i,t}} + \textbf{Fixed Effects}$$

where $ESG\_Fund_{i,t}$ is the proportion of ESG funds in the investment menu of plan i in year t. Following the approach of Pool, Sialm and Stefanescu (2016) and Tran and Wang (2023), the control variables include certain plan characteristics that are unrelated to participant attributes but may impact option choices. To account for across-plan variations and temporal trends in ESG preferences and menu changes, we introduce both plan and year fixed effects.

The coefficient β in the first column of Table 1.3 shows that with 100 more scandals happening nearby in the previous year, the plan would have 0.5 percentage points more ESG funds in the 401(k) investment menu. To study the dynamic reason for scandals triggering more ESG funds, we identify four types of menu changes: adding, deleting, and replacing funds in and out. The dependent variables are adjusted to the different types of menu changes while keeping the control variables constant in the following columns. In column 2 of Table 1.3, we observe that, conditional on plans being added at least one new fund, ESG scandals significantly increase the likelihood of adding ESG funds to the menu. An additional 100 scandals is associated with a 1.4 percentage point increase in adding an ESG fund to the menu. Columns 3 to 5 show that the coefficients for all other types of menu changes are statistically insignificant. This suggests that investors predominantly react to ESG scandals by incorporating new ESG funds into their investment menu. We focus on the number of newly added ESG funds in the menu in the following context, as it holds potentially meaningful causal implications.

In the subsequent columns of Table 1.3, we employ logistic regression models to assess the robustness of the relationship between ESG scandals and the likelihood of adding ESG funds to 401(k) menus. The logistic regression model is formulated as follows:

$$\text{Prob}(ESG\_Add_{i,t} > 0) = \Lambda(\beta Scandal_{i,t-1} + \boldsymbol{\Gamma Controls_{i,t}} + \textbf{Fixed Effects})$$

Here, Λ represents the logistic cumulative distribution function. Year fixed effects are also included. Given the "perfect separation" problem in the logit model when there is no within-group variation of the dependent variable, we introduce geographic fixed effects for the first two digits of the zip code in the unconditional case. Because changing the menu could be influenced by various endogenous factors, the results may vary depending on whether we treat ESG fund addition as a single behavior or as a behavior based on menu changes. We differentiate between unconditional and conditional ESG fund additions. In the conditional case, we treat ESG fund addition as a discrete behavior, assigning zero for plans that have menu changes but have not incorporated ESG funds into their menus within a given year. For this case, the sample only contains plans that have menu changes in a given year. In the unconditional case, we assign zero for all plan-year observations without ESG funds added. The unconditional case has all observations without any constraints.

Column 6 of Table 1.3 illustrates that ESG scandals from the previous year significantly contribute to explaining the variation in adding ESG funds, under the condition that the plan decides to alter its menu. In this "Change" conditional scenario, where a plan adds or deletes one or more funds from its menu, a 100-scandal increase leads to a 0.13 percentage point rise in the probability of adding an ESG fund. Similarly, column 7 demonstrates that, in unconditional cases, a 100-scandal increment is linked to a 0.24 percentage point increase in the likelihood of ESG fund addition. The marginal effect is even higher than the conditional case when using a different fixed effect than in column 6. Column 8 verifies the persistence of these results with the different fixed effects from those in column 6.

1.3.3 Mechanism of Scandal Impacts on Portfolio Changes

The portfolio changes after the scandals could be attributed to different reasons. From a portfolio choice theory, the shocks could hit the pecuniary part, including expected returns, systematic risks, and risk aversions. It could also affect the non-pecuniary part, i.e. the sustainable preferences. With the help of RepRisk, we focus on the three main dimensions of ESG scandals: severity, reach, and novelty. Severity measures how bad the scandal is regarding social impact, which has a value ranging from 1 to 3. Reach measures how much scandals are publicly known for everyone at the story date, ranging from 1 to 3. Novelty is whether the scandal is a new issue to the firm or some related issues have already happened to the firm. We use HS, which is the number of highest severe scandals, HR, which is the number of scandals known to all investors at the story day, and New, which is the number of novel scandals, to substitute the $ESG\_Scandal$ as the key independent variable. Columns 1 and 4 of Table 1.5 show that high-severity scandals could not explain the variation in ESG investment. However, the occurrence of high-reach and new scandals contribute to the variation significantly, as shown in the other columns. The psychological part is more essential than the real-impact part in inducing ESG investment decisions. On the one hand, there are many other ways to hedge the severe ESG concerns, for example, moving away from this area. On the other hand, scandals act as social categorizations to navigate investors' latent preferences. [25]

1.3.4 Heterogeneous Investor Features in ESG Fund Addition

We proceed to explore potential heterogeneity with respect to investor demographic features in response to ESG scandals. This part involves segregating all plan-year observations into subsamples defined by investor characteristics. We still follow the two conditions above in the fund-addition study, focusing on all plans or just those with menu changes.

In Panel (a) of Table 1.6,we employ both a logit model and a continuous dependent variable panel OLS to examine the likelihood of ESG funds being added to menus in Democratic and Republican areas. The regression models are identical to Table 1.3. In columns 1 and 3, the results highlight that plans in Democratic areas respond to ESG scandals with a higher probability of ESG fund addition. The marginal effect of ESG scandals on this likelihood stands at 0.3 percentage points per 100 incidents. Conversely, columns 2 and 4 reveal that plans situated in Republican areas yield insignificant propensities. These outcomes persist robustly when using an OLS framework and employing other fixed effect specifications, as evidenced in the final four columns. Supporters of the Democratic Party consistently endorse ESG considerations. A surge in ESG scandals triggers them to include sustainable funds in their portfolio, possibly as a hedge against scandal risk. However, adherents of the Republican Party might perceive ESG as an unsubstantiated concept or may accord it less importance within their investment strategy. Thus, ESG scandals might merely register as noise, evoking no discernible response in terms of ESG investment. Alternatively, the pattern could stem from right-leaning individuals valuing real-world outcomes but disbelieving in ESG-focused investments. While the dataset does not permit a definitive distinction between these hypotheses, both

[25] We show more about the robustness of the mechanism and the scandal assumptions in the appendix. Unexpected scandals are more likely to cause menu changes as well, which also partially supports the mechanism of changing latent heterophily to explicit choices.

scenarios ultimately converge on the same conclusion.

In the first four columns of Panel (b) of Table 1.6, we divide the plans into two subgroups—young and old—based on the flow-based age measure [26]. We use the median value in each year to get the subsamples and annually rebalance them. The regression settings are the same as in the first four columns of Panel (a). Columns 1 and 3 indicate that the old group exhibits coefficients of greater magnitude and higher levels of significance than the young group, as shown in columns 2 and 4. In the unconditional context, both young and old groups respond to ESG scandals by adding ESG funds. However, the magnitude of the marginal effect for old investors is nearly twice as large as for the young group. This suggests that ESG scandals serve as a reminder or impetus for old investors to prioritize ESG within their portfolios, compensating for previous underreaction as shown in Table 1.2. To shed light on the reasons behind this, we delve into fund flow in the subsequent sections. Similarly, when using the TDF year as the age measure and the same rebalancing method in Panel (b), a similar trend is observed, with the old group demonstrating a greater propensity to integrate ESG funds into their menus. While the conditional context (columns 7 and 8) fails to yield significant coefficients, the unconditional case underscores that the marginal effect from ESG scandals is still twice as pronounced for the old group than for the younger investors.

In Panel (c) of Table 1.6, we continue the approach by focusing on subsample regressions based on the welfare proxies. We follow the same subsample determination as in Panel (b), which uses the median value of the measure each year to split and rebalance the entire sample annually. In the first four columns, where the deferral amount serves as a proxy for wealth, both rich and poor investors exhibit an increased inclination to add ESG funds to their menus after the occurrence of more ESG scandals. The difference in marginal effects is not statistically significant. When the account balance is used as the proxy for wealth, as seen in the last four columns, solely the rich investors manifest an experience effect on ESG fund addition. The magnitude aligns closely with the rich group's deferral measure. In contrast, poor investors do not exhibit a corresponding trend of adding more ESG funds.

[26] i.e. $Flow_{TDF \leq 2040} / Flow_{TDF > 2040}$

The distinctions between the two proxies are essential. Deferral reflects intratemporal constraints, thus indicating that individuals with lower incomes and tighter financial limitations still exhibit a menu reaction to ESG scandals. Account balance, conversely, is influenced by age and wealth accumulation. Combining these insights, it becomes evident that individuals with greater wealth consistently respond to ESG scandals by incorporating ESG funds into their investment menus. On the other hand, individuals constrained by intratemporal considerations still exhibit intense reactions, possibly due to their sensitivity to specific types of scandals that affluent individuals might not find as important. However, when faced with low accumulated wealth, individuals may not invest in ESG even in the face of frequent scandals due to pecuniary concerns. This suggests that the effects of ESG scandals on investment choices are complex and multifaceted, shaped by a combination of financial, age-related, and psychological factors.

We proceed to test whether investors' beliefs, specifically their political leanings, are prior to other features. In Table 1.7, the regression coefficients β are reported based on the model in columns 7 and 8 of Table 1.3, using a double-sorted subsample based on political leanings and two other factors. We split the plans in each political area into two groups based on the age or the wealth measure and annually rebalance it.

In the upper panel of Table 1.7, we examine the unconditional case. In Republican areas, regardless of age or wealth, the β coefficients consistently show insignificance. Individuals with conservative political leanings, even old or wealthy, appear to be insensitive to ESG scandals in terms of extensive margins. The Democratic subsamples demonstrate a consistent pattern that mirrors Table 1.6. Older and wealthier investors within the Democratic subsamples are more inclined to include ESG options when confronted with ESG scandals. Remarkably, even financially constrained individuals (as measured by deferral) within Democratic areas exhibit a significant coefficient. Moving to the conditional case in the lower panel, the results are consistent. Notably, wealthy individuals in Republican areas exhibit a marginal significance in adding ESG funds to their menus due to ESG scandals. However, older investors using the TDF-based measure do not show a propensity to add ESG funds, which aligns with the discussion in columns 7 and 8 of Table

1.6.

Overall, these findings suggest that political beliefs have explanatory power that is orthogonal to responses to ESG scandals through fund additions. Investors should embrace ESG concepts before investing in them. Democratic-leaning investors tend to prioritize ESG options in response to scandal shocks. However, personal characteristics such as financial constraints, age, and social experiences also play a significant role. Financially constrained individuals exhibit unexpected resilience in adopting ESG funds, while older and wealthier investors exhibit a mix of financial capacity and ESG concerns.

1.4 Categorizing ESG Scandals and Deciphering Investor Responses

In this section, we categorize ESG scandals into distinct groups and examine how the addition of ESG options varies across these categories. Our goal is to pinpoint the specific types of scandals that trigger each investor group to include ESG options in their portfolios and compare the magnitude of their effects.

1.4.1 Three Main Categories of Scandals: Environment, Social, and Governance

We begin by categorizing each scandal into the three fundamental aspects of ESG—environmental, social, and governance scandals—and comparing the roles of these three aspects in ESG investment decisions. Given that a single scandal may fall under multiple categories, we resolve multicollinearity by treating the occurrence of each type of scandal independently in columns 1–3 and columns 5–7 of Table 1.8. For scandals that can only be attributed to a single category, we form their occurrence as a vector of independent variables in columns 4 and 8. The regression model used is consistent with that presented in columns 7 and 8 of Table 1.6, with the only difference being the replacement of $ESG\_Scandal_{i,t-1}$ with the occurrence of separate categories.

The results in columns 1 and 5 demonstrate that social scandals have a significant positive impact on the likelihood of adding ESG funds to the 401(k) menu, both unconditionally and conditionally. A 100-unit increase in social scandals nearby leads to a 0.35% increase in the probability of adding ESG funds unconditionally, and a 1.1% increase conditionally. These effects are notably larger compared to those observed in Table 1.3. However, columns 2 and 6 reveal that environmental scandals show no significant correlation with the addition of ESG funds. Similarly, governance scandals do not exhibit a substantial influence.

Overall, the findings suggest that only social scandals significantly prompt investors to add ESG options to their menu. These results remain robust when using the vector of all three types as the independent variables, as seen in columns 4 and 8. Social scandals seem to have a more direct impact on investor attention and daily life, resulting in a significant trigger on extensive decisions. This is different from some previous studies[27] that have emphasized the impact of environmental aspects, such as carbon emissions.

Social scandals are associated with social conflicts and may lead to high financing costs for the affected companies, which act as cognitive shocks to local residents. Such scandals are always shocking enough and less likely to be predicted. While the environmental aspect should theoretically have a similar effect, the absence of significance could imply that investors have a low willingness to pay for extensive margins on environmental issues. This could result from reacting to environmental scandals in other efficient ways. Facing air pollution issues, the most efficient way is to move to a clean area or invest in health insurance rather than ESG funds. In addition, environment scandals usually lack novelty compared with social ones.

1.4.2 Mapping Heterogeneity to ESG Scandal Categories

We move to comparing the roles of different scandals in heterogeneous investor preferences. To have a clear cross-matching framework across types of scandals, investor characteristics, and ESG decisions, we use the same subsample classifica-

[27]For example, Bernstein, Gustafson and Lewis (2019) highlight natural disasters, and Bolton and Kacperczyk (2021) show the importance of carbon risk.

tions as in Table 1.6. Table 1.9 has the results, with each pair of coefficients (β) and standard errors corresponding to a distinct regression.

In the upper panel, we consider the unconditional case that treats ESG fund addition and menu change as a single behavior. The findings from the first column highlight a significant effect of social scandals. Except for the Republican-leaning group, investors across all other categories—regardless of age or wealth—demonstrate significant inclinations to ESG fund addition to their 401(k) menus. Notably, older investors display a marginal effect that is 0.2 percentage points higher per 100 social scandals compared to their younger investors. The magnitude of reaction does not seem to significantly differentiate between affluent and less affluent investors.[28] Political leaning still plays a pivotal role, as only Democratic supporters consistently respond to social scandals. The wealth of life experiences of older individuals makes them more sensitive to social issues, resulting in a heightened extensive response. However, young investors might lack a comprehensive understanding of social rights, and their limited financial capacity could constrain additional investments.[29] Environmental and governance scandals appear to have no significant impact on any group of investors. An interesting outlier emerges: older investors, as measured by the TDF flow, exhibit a significant inclination to add ESG funds in response to governance scandals. This is unsurprising, considering that governance scandals are closely correlated with social ones, both conceptually and statistically. Old investors, who care more about the governance part of the social scandals, in this way show significant results.

In the lower panel of Table 1.9, we shift the focus to the conditional case where all observations are plans with menu changes. Democratic voters exhibit an interest not only in social scandals but also in environmental ones in this scenario. Investors, particularly those in Democratic-leaning areas, demonstrate concern for environmental scandals; if given the opportunity to change the menu, they opt to do so. It is important to note that the decision-making process behind menu changes is intricate and influenced by various factors. Therefore, the results are not in contradiction with those of the unconditional case. Nonetheless, it is evident

[28] Although the magnitude of the coefficients seems different, the comparison t-stat, 1.27, is insignificant.

[29] They may have already had the ESG options according to Table 1.2.

that environmental scandals may not carry the same significance for all investors. Overall, we observe a stronger marginal effect for environmental scandals in comparison to social ones.

Older investors, as measured by TDF flows, continue to respond to social and governance scandals, exhibiting a statistically significant marginal effect where the propensity of adding ESG funds to the menu increases by 1% per 100 scandals. In contrast, the younger group shows a higher inclination to add ESG funds in response to environmental scandals, with a magnitude nearly twice that of their response to social scandals. This heterogeneity in responses to different types of scandals by young and old investors can be attributed to life cycle concerns. Environmental issues such as pollution and rising sea levels have a profound impact on the future, affecting the younger generation more significantly due to their longer expected lifespan. Conversely, the older group, especially those without altruistic tendencies, may dismiss these concerns as irrelevant given their shorter remaining lifespan. Social scandals, which have a more direct impact on daily life, are more relevant to older investors, who are more likely to change preferences in response to these. The similar trend is observed when measured by TDF age in Table 1.9, where young investors show a higher level of responses to social scandals, but they still care about environmental ones.[30]

Considering the subsamples based on wealth, affluent investors in both measures add ESG funds in response to social scandals. They are not reacting to environmental scandals in this conditional case. Poor investors show no responses to any type of scandal except when account balance is used as a measure. This can be attributed to the correlation between account balance and employee age. As the account balance overlooks wealth accumulation effects, younger employees typically possess lower account balances. If young investors from new plans are deeply concerned about the environment and have the opportunity to add a new fund to their menus, this significant coefficient is unsurprising.

In summary, the social component of the ESG index has emerged as the most influential factor driving ESG investments. Investors leaning towards the Demo-

[30]The difference between these two coefficients is marginal. By using the single attributed scandals as shown in Table 1.8, the difference has a t-stat of 1.32.

cratic Party exhibit heightened sensitivity to social scandals. Older and wealthier individuals are more prone to react to social scandals, while younger investors also respond to environmental incidents. The governance issues have no impact on ESG fund investment.

1.4.3 Specific Scandal Issues

To gain a more granular understanding of how different types of scandals influence ESG fund addition, we have classified scandals into 13 distinct categories. These categories are not mutually exclusive, as a single scandal can fall under multiple categories. One issue could be attributed to more than one of the three basic dimensions. For instance, categories like *Legal & Compliance* and *Political & Governance* are classified as both social and governance issues. In Table 1.10a, we present the results of the same logistic regression model, using the occurrence count[31] of each scandal category as an independent variable. Each cell in the table contains the coefficient (β) and standard error resulting from a single regression for the corresponding subsample group indicated in the first column.

The first row of Table 1.10a reveals that scandals falling within categories such as *Cyber & Privacy Concerns*, *Labor & Employment*, *Legal & Compliance*, *Political & Governance*, and *Social & Human Rights* are pivotal triggers for the inclusion of ESG funds. *Political & Governance* exerts the greatest impact, with a marginal effect of 5 basis points for each additional scandal in this category. Notably, these categories are all linked to social aspects. The significance is unsurprising as these issues directly influence real-life provisions and act as attention shocks. *Legal & Compliance* holds the second-highest marginal effect and is related to the attention in provisions. *Cyber & Privacy Concerns* with the highest mean value across all valid issues are also highly related to real-life rights and unpredictable shocks. This issue is scarcely addressed as an ESG investment trigger in prior studies. Indeed, the *Cyber & Privacy Concerns* category has a mean value of 7 per year within each area, as evidenced by Table 3.7b, underscoring its notable importance. The impact of cybersecurity and data privacy concerns on individuals' lives becomes evident

[31] Without normalizing to 100.

only when investors personally experience such incidents, subsequently motivating their engagement in ESG investing.

Moving on to the subsamples based on investor characteristics, it becomes evident that *Cyber & Privacy Concerns* do not elicit significant reactions for individuals with lower income or Republican leanings. The lack of response from individuals with lower income can be attributed to the fact that cybersecurity concerns are less directly related to their wealth and, therefore, less likely a shock to them. On the other hand, Republican voters, often more conservative in their investment outlook, might be less inclined to embrace ESG investments. Only *Political & Governance* issues register as a significant concern for Republicans. This association between conservative beliefs and preferences for *Political & Governance* issues aligns with certain stereotypes.

Shifting focus to *Labor & Employment*, it appears that only older and wealthier individuals show a notable concern. In labor concerns, young individuals may be more engaged in advocating for labor rights through direct actions rather than just investment choices. Conversely, older and wealthier individuals might be motivated by both their experiences and greater financial capacity. Examining the *Social & Human Rights* category, we find that all characteristics except for Republican voters exhibit reactions. Notably, individuals with lower income demonstrate the strongest response, while wealthier individuals display the weakest. This trend can be attributed to the fact that those with lower incomes are more likely to empathize with those impacted by labor abuse. They may perceive themselves as victims of such issues. The scandals evoke their incentive to invest into ESG.

For categories that show insignificance for the entire sample, we find that certain issues become effective for specific groups. In the first column, concerns related to animal welfare and biodiversity appear to resonate with Democratic Party supporters. This could be attributed to the heightened animal protection awareness prevalent among left-wing investors. Although the magnitude of impact is considerable, the unconditional mean value for such scandals is low.[32] In the case of *Technology & Surveillance*, despite its lack of significance in the entire sample, it exhibits an impact on older and wealthier individuals. It is plausible that

[32] Only 1.8 according to Table 3.7b.

only areas with significant exposure to technology-related concerns are affected by such scandals. For instance, a farmer in Tennessee may not have the same level of concern about technology scandals as would an IT engineer in Silicon Valley. Similarly, the *Weapons & Security* category only appears to have an impact on younger and less affluent individuals. This result aligns with the hypothesis that those who are more likely to be directly impacted by weapons and security issues are more inclined to respond to these types of scandals.

Table 1.10b presents the coefficients of the same regressions as shown in Table 1.10a, but observations are conditioned on plans that have changed their menus. In the overall sample, the significance of *Cyber & Privacy Concerns*, *Legal & Compliance*, and *Social & Human Rights* remains, with even larger magnitudes. Intriguingly, *Animal & Biodiversity* becomes significant in the whole sample, with an increase in propensity of 1.5 basis points due to each additional scandal in this category. This effect is primarily driven by Democratic-leaning plans and young investors. When looking at deferral-based measurements, affluent individuals demonstrate more concern, while account balance-based measurements show that poorer individuals are more responsive. This could be reconciled by high-income yet young investors who may react to these scandals.

Another notable observation is the increased significance of *Substance & Social Ethics*, an issue that previously had no effect in the unconditional case. All subsamples exhibit a concern for this issue. This finding could suggest that plans with a blend of older employees[33] and a higher influx of younger employees[34] are sensitive to this issue. *Substance & Social Ethics* addresses societal issues such as alcohol, gambling, or pornography, making it more relevant to young, affluent individuals with Democratic leanings who value it through their ESG investments. This issue has a mean occurrence rate of 1.4 per area per year, and the corresponding marginal effect is 0.13 basis points. To explain the difference between unconditional and conditional fund addition due to this issue, individuals might perceive real-world activities as a more suitable response than mere ESG investment, but when the opportunity arises to add ESG funds to their menus, it seems

[33] Measured by TDF Age.
[34] Measured by the Flow Age.

to spur them into action.

What was previously significant in the whole sample in the unconditional case now becomes insignificant, including *Labor & Employment* and *Political & Governance*. When we focus solely on plans with changing menus, which are in an unstable state, the influence of *Labor & Employment* on the decision to add an ESG fund may be overshadowed by other factors such as the labor movement or menu modifications. Similarly, *Political & Governance*, which was the most significant issue in the unconditional case, remains relevant for Republican-leaning investors in the conditional case. Yet, for other investor characteristics, its impact diminishes. This could be due to a similar reason as with *Labor & Employment*. In comparison to all active plans, these scandals seem to have no effect on the decision to add ESG funds.

In the conditional case, *Environmental Issues* have an impact on the decision to add ESG options, but this impact is observed only within Democratic-leaning areas. Encouraging individuals to react solely to environmental issues is indeed a challenge. As argued in various literature, firms that are more exposed to environmental risks are often the ones allocating more resources to ESG efforts (Cohen, Gurun and Nguyen, 2020). Consequently, environmental scandals may not carry significant weight in extensive margins, as investors could anticipate that companies will take measures to address such issues. Those who do respond to these scandals, such as Democratic voters, might be reinforcing their existing beliefs or adopting a more pessimistic outlook.

The two panels in Table 1.10 clearly demonstrate that specific incidents, including some that traditional ESG metrics might have overlooked, can significantly influence the likelihood of adding ESG funds to the menu. Investors with varying characteristics may have distinct levels of exposure to different scandals, which in turn affects their heterogeneous reactions. These findings highlight the importance of offering a diverse range of ESG investment options for different investors. Comparing across ESG elements, social aspects would be more likely to trigger ESG fund incorporation, while environmental aspects have a narrow triggering effect on a certain group of investors.

1.5 Heterogeneity in ESG Fund Flows: Investor Responses to ESG Scandals

Shifting the focus to the existing ESG fund flows as the decisions, we apply the same logic as in Sections 3 and 4, revealing heterogeneity and comparing the ESG elements. Unlike menu changes, these fund flows only shed light on the existing ESG funds in the menu. Analyzing fund flows allows us to move from portfolio choices to portfolio weights, assessing the extent to which investors increase their investments in response to various ESG scandals.

1.5.1 Employee Fund Flows and ESG Scandals

We first examine the impact of aggregate ESG scandals on ESG and non-ESG funds separately, verifying that these incidents lead to increased investment in ESG funds while not significantly affecting non-ESG funds. In Table 1.11, we employ a panel OLS regression model:

$$Flow\_Employee_{i,j,t} = \beta Scandal_{i,t-1} + \mathbf{\Gamma} \boldsymbol{Controls_{i,j,t}} + \textbf{Fixed Effects}.$$

This model is applied to each fund j within plan i during year t. The employee flow for each fund is computed using the formula:

$$Flow\_Employee_{i,j,t} = \frac{V_{j,i,t} - V_{j,i,t-1}R_{j,t}}{A_{i,t-1}}.$$

Here, $V_{j,i,t}$ represents the investment value of fund j within plan i at the end of year t, $A_{i,t-1}$ is value of the total assets in plan i at the end of year $t-1$, and $R_{j,t}$ denotes the net return of the same fund from time $t-1$ to t. This calculation only applies to ongoing funds j without replacements within the given year t. In the case of funds added or deleted during year t, we consider all the money to be taken in and out. However, for replacement funds, if fund j is brought in to replace fund k within the menu, the fund flow is calculated using the formula:

$$Flow\_Employee_{i,j,t} = \frac{V_{j,i,t} - V_{k,i,t-1}R_{k,t}}{A_{i,t-1}}.$$

In this scenario, it is assumed that all assets originally held in fund k are transferred to fund j. While this assumption might not be highly stringent, it captures the inactivity of participants when replacements occur. However, it is important to acknowledge that if the replacement fund does not align with the participants' asset management preferences, then this measure is biased.[35] All employee fund flows are standardized to percentage levels and winsorized at 1% and 99%.

We extend control variables $\boldsymbol{Controls_{i,j,t}}$ to ensure that they have the potential influence on fund flows within each plan. Specifically, we keep two plan-level control variables: the total asset value of the plan and the number of investment choices in the plan. Funds in large plans with many other options inherently experience low flows per year. Additionally, we include the three-year accumulated net returns ($Ret_{t-1,t-3}$) and three-year monthly volatility ($Vol_{t-1,t-3}$). The accumulated net returns account for the fund's performance over a three-year horizon, striking a balance between short- and long-term concerns. All regressions control for year- and plan-level fixed effects.

The first column in Table 1.11 presents a significant relationship between ESG scandals and the flow to ESG-focused funds. Specifically, for every 100 additional ESG scandals experienced, the flow to ESG funds increased by 4.2 basis points. When considering non-ESG funds, a contrasting pattern emerges. In the second column, an increase of 100 scandals is associated with a decrease of 1.8 basis points in the flow to non-ESG funds. An inflow to ESG funds naturally corresponds to an outflow to non-ESG funds. This provides evidence that ESG scandals not only trigger new fund additions but also an inflow to existing ESG funds. The following analysis narrows the focus to ESG funds in the 401(k) sample.

Similar to the analysis on fund additions, we perform consistent regressions across the subsamples based on demographic characteristics. A notable outcome emerges in relation to political leaning, as shown in columns 3 and 4. Unlike Democratic-leaning party voters, who show no significant increase in fund inflow to ESG funds following ESG scandals, Republican-leaning investors exhibit higher inflows. To explain this difference in heterogeneity between fund addition and fund flow, several factors come into play. First, Democratic-leaning investors who have already added

[35]The replacement happens with a low probability of 13% for all funds and 9% for ESG funds.

ESG funds to their portfolios might perceive ESG-related scandals as relatively unsurprising. This perception could be attributed to their willingness to pay for ESG factors, as explored in Heeb et al. (2023). Second, the inflow from Republican-leaning investors could be attributed to the heresy nature of plans with ESG funds in Republican-leaning areas. The presence of ESG funds in conservative regions indicates a strong belief in the value of ESG, possibly even more pronounced than in Democratic-leaning areas. This belief may also be connected to the context of political swing areas. Republican-leaning areas with plans featuring ESG funds exhibit a higher rate of swing toward Democratic-leaning areas. Specifically, if a Republican-leaning area has at least one plan with ESG funds in the menu, the swing rate toward Democratic-leaning areas is 1.17%. In contrast, areas without such plans exhibit a transformation rate of only 0.32%. This swing propensity, coupled with the catalytic role of ESG scandal shocks, contributes to the inflow to ESG funds. While this coupling effect is not strong enough to cause changes in menus, it does impact the flow to existing ESG funds. The number of plans with ESG funds in the Republican areas is considerable. Although plans in the Republican areas have a relatively lower probability of adding ESG funds, the total number of plans in Republican areas is greater than that in the Democratic areas. This implies the fact that every ESG investor matters. If some funds want to target the ESG heresy in Republican areas, they still have a promising market.

The examination of wealth is consistent with the findings from the fund-addition analysis. Wealthier investors demonstrate a greater willingness to increase investments in ESG funds in response to ESG scandals. This observation is intuitive, as individuals with higher deferrals likely possess greater financial capacity to invest without being hampered by pecuniary constraints. Interestingly, the account balance measure does not exhibit significant variations within each subsample. This lack of variation could similarly be attributed to the influence of political factors and beliefs in shaping investment decisions.

To examine the role of investor political beliefs and the heterogeneity across different age and wealth groups in ESG fund flow, we introduce the same subsample division as presented in Table 1.7. As shown in the fund-addition case, political leaning is found to be prior to the other two factors. We aim to test this with respect to ESG fund flows in Table 1.12. In the left part of the table, plans in Republican-leaning areas exhibit some significant coefficients. Young and wealthy investors within this segment are shown to channel greater inflows into ESG funds in response to ESG scandals. Notably, the marginal effect is particularly significant for young investors, reaching an impressive 7 basis points per 100 scandals. This time, account balance and TDF age are significant characteristics causing heterogeneity. No single subsample of Democratic-leaning investors manifests a statistically significant coefficient. Clearly, in ESG investment decisions, beliefs are still orthogonal to wealth and age. Investors should possess a certain level of willingness to pay for ESG, whether extensively or intensively, in order to react to ESG scandals.

1.5.2 Specific Scandal Issues: Flow-Based Analysis

We undertake the same categorization of the ESG scandals in Section 4 to compare the roles of different scandals in triggering ESG fund flows. We also aim to quantitatively study the inflow induced by each type of scandal and check whether investors who are not sensitive to aggregated scandals would react to certain categories of scandals in fund flows.

In the initial three columns of Table 1.13, we disaggregate the scandals into the traditional categories of social, environmental, and governance issues. Using the occurrence of each scandal type as the independent variable, the structure of the table mirrors that of Table 1.10. The first row of the table showcases the impact of social and environmental scandals on ESG fund inflows within the entire sample. Both environmental and social scandals influence fund inflows, with environmental ones exerting an effect twice as strong as that of social scandals. This departure from the fund-addition case suggests a complex interplay between investor reactions and scandal categories. Notably, the reaction remains largely confined to Republican-leaning investors, who are particularly responsive to both social and environmental scandals. Younger and wealthier investors react prominently to environmental incidents, underscoring their heightened sensitivity to environmental concerns. Governance-related incidents show no significant effect in all cases.

The subsequent 13 columns transition to a more detailed analysis of the 13 specific issue categories. Starting with the *Animal & Biodiversity* category, it is a concern primarily among wealthy individuals. In general, animal welfare and biodiversity might have little attention or social impact. *Cyber & Privacy Concerns*, which showed importance in ESG fund additions, now plays a diminished role in fund inflows. The significance of this category is now primarily observed among older and wealthier investors. The earlier prominence of this issue in the fund-addition context could be linked to investors' initial assessments in shaping their ESG investment but without a high willingness to pay more.

The *Economic & Consumer Rights* category becomes a significant shock to ESG fund inflows. A single occurrence of such a scandal leads to an increase of 0.4 basis points in ESG fund inflows. This effect holds true for both Democratic and Republican areas. Scandals in this category are related to real-world economic issues, such as tax evasion or unethical business practices. Despite the challenge of persuading investors to add ESG funds to their menus, the results suggest that they, especially those who are young and rich, react by higher portfolio weights given that they already hold ESG funds.

Furthermore, the importance of environmental concerns in driving ESG fund inflows is reaffirmed by the significance of the *Environmental Issues* category. A one-standard-deviation increase in *Environmental Issues*[36] leads to a 3.1 basis point rise in ESG fund inflows. This effect is not exclusive to young investors; even older investors, as measured by TDF age, exhibit a response. Additionally, both poor and rich groups actively invest more in ESG funds in response to this category. While the social aspect of scandals might prompt ESG fund addition, the environmental dimension triggers them to allocate more resources. A similar dynamic is observed in the *Health & Safety* category, which includes issues like epidemics and pandemics. Although the effect size is relatively small within the sample, it is plausible that had our study period included the COVID-19 pandemic, the magnitude of the effect could have been more substantial. This underscores the potential impact of health-related concerns on investor behavior in the context of ESG investment.

[36] Equivalent to 31 additional occurrences, as shown in Table 3.7b.

The *Labor & Employment*, *Legal & Compliance*, and *Political & Governance* categories in Table 1.13 show no significance in the fund-flow case. These categories act as attention and real-life shocks to fund additions but are not harsh enough to drive further inflows when there are ESG options on the menu. Investors who care about these issues just have the willingness to include ESG funds in their portfolios when they do not already have those options in their portfolios. On the other hand, those who react to general scandals by ESG fund inflow are not sensitive to these issues. These observations emphasize that the impact of different types of scandals on investor behavior is multifaceted. Some social issues cause a group of investors to add ESG funds to their portfolio, but these issues do not easily induce more weight on ESG funds.

The significance of *Social & Human Rights* issues in driving ESG fund inflows remains notable. The influence of social scandals on ESG inflows is particularly evident in Republican-leaning areas and among young investors. This effect, ranging from 1.5 to 3 basis points per additional standard deviation of social scandals, underscores the enduring importance of human rights concerns within conservative spheres.[37]

These findings suggest that when ESG funds are in the portfolios, environmental issues play an equivalent role in triggering fund flows, especially for young and heretical investors. Social issues, especially in economic development and human rights, are still important. To understand this, scandals with temporal life impacts and direct attention attraction are more likely to cause ESG fund addition; scandals with long-term life impacts and controversial attention attraction cause more ESG fund flow. The diverse reactions to different scandals highlight the heterogeneity of investor concerns. The difference between the results of fund addition and fund flow analyses also prompts us to delve into the underlying factors that guide investor reactions to ESG scandals.

[37] As mentioned in the previous section, the political swing rates are high in these areas.

1.5.3 Sample-Selection Bias and the Heckman Model

Concerning ESG fund flows, a significant sample-selection issue has arisen: ESG fund flow observations were restricted to plans that already included ESG funds. This inherently introduces bias when evaluating the unconditional impact of ESG scandals on ESG fund flows. We employ the approach by Heckman (1979) to mitigate the sample-selection bias and provide an unbiased analysis of the influence of ESG scandals on ESG fund flows. This also helps us understand the attention effect of ESG scandals.

To elaborate further, denote an unobserved variable as $D_{i,t}$, representing the propensity of having an ESG fund in the investment menu. This variable is determined by:

$$\boldsymbol{D} = \boldsymbol{z}\gamma + \boldsymbol{\epsilon}.$$

where $\boldsymbol{z}$ is a set of plan-specific characteristics. We could only observe ESG fund flow $D^*_{i,t} = 1$ when $D_{i,t} > 0$, otherwise $D^*_{i,t} = 0$. The probability of ESG fund flow $Flow\_Employee_{i,j,t}$ could be observed with the probability

$$Prob(\epsilon_{i,t} > -\boldsymbol{z_{i,t}}\gamma) = 1 - \Phi(-\boldsymbol{z_{i,t}}\gamma) = \Phi(\boldsymbol{z_{i,t}}\gamma)$$

if we assume that $\boldsymbol{\epsilon} \sim \boldsymbol{N(0, I)}$. Based on the model to explain the scandal's effect on ESG fund flow, the conditional mean of ESG fund flow should be

$$\begin{aligned}
\mathbf{E}[Flow_Employee_{i,j,t}|D^*_{i,t} = 1] &= \beta Scandal_{i,t-1} + \boldsymbol{\Gamma}\, \boldsymbol{Controls_{i,j,t}} + \\
&\quad \textbf{Fixed Effects} + \mathbf{E}[u_{i,j,t}|D^*_{i,t} = 1] \\
&= \boldsymbol{X\beta} + \mathbf{E}[u_{i,j,t}|\epsilon_{i,t} > -\boldsymbol{z_{i,t}}\gamma] \\
&= \boldsymbol{X\beta} + \rho\sigma_u \underbrace{\frac{\phi(-\boldsymbol{z_{i,t}}\gamma)}{\Phi(-\boldsymbol{z_{i,t}}\gamma)}}_{\text{Heckman's } \lambda}.
\end{aligned}$$

If the errors in the fund flow and fund selection models are uncorrelated ($\rho = 0$), we can safely apply ordinary least squares to uncover unbiased estimates for $\boldsymbol{\beta}$ and we can neglect the selection equation part of the model. However, in our study, unobservable factors related to plan participants and sponsors influence both error terms. For instance, if a sponsor strongly supports ESG principles, they might encourage participants to add ESG funds to the menu and stimulate greater investment in them. To estimate the unconditional impact of ESG scandals on employee fund flows, we follow these steps outlined in Heckman (1979):

- Run the probit on the selection model

$$Prob(D_{i,t} > 0) = \Lambda(\gamma \boldsymbol{z_{i,t}}) = \Lambda(\beta_0 Scandal_{i,t-1} + \boldsymbol{\Gamma_0} \boldsymbol{Controls_{i,t}})$$

.

- Recover the estimated Heckman's λ ($\frac{\phi(-\boldsymbol{z_{i,t}}\gamma)}{\Phi(-\boldsymbol{z_{i,t}}\gamma)}$).
- Using OLS, run the regression

$$\begin{aligned} Flow\_Employee_{i,j,t} = \quad & \beta Scandal_{i,t-1} + \boldsymbol{\Gamma_1} \boldsymbol{Controls_{i,j,t}} + \\ & \rho\sigma_u \frac{\phi(-\boldsymbol{z_{i,t}}\gamma)}{\Phi(-\boldsymbol{z_{i,t}}\gamma)} + \textbf{Fixed Effects}, \end{aligned}$$

  where $\rho\sigma_u$ is treated as a single parameter to be estimated.
- Adjust standard errors to account for the fact that Heckman's λ is an estimated covariate in the above model.

Table 1.14 presents the results of the above procedures for the entire ESG fund sample as well as for subsamples based on the five variables. In the lower panel, the variable D represents the estimated coefficients of β_0 from the first-step probit model. We also incorporate all other control variables that are consistent with the fund-addition part and largely associated with the selection of ESG funds. The results are notably consistent with the earlier estimations in Section 3.

In the upper panel of Table 1.14, we introduce Heckman's λ as an additional regressor, which provides insights into both the direct and indirect effects of scandals. Specifically, λ signifies the influence of having an ESG fund in the menu. The results clearly indicate that only λ coefficients are statistically significant. Surprisingly, the scandals themselves do not appear to affect ESG fund flows directly. Conversely, the high significance of λ in most of the subsamples underscores the importance of having ESG funds in the menu. This finding aligns with previous conclusions from the conditional flow tests. The significant negative coefficients of $\rho\sigma_u$ imply that $\rho < 0$, indicating a negative correlation between factors influencing fund addition and fund flow. This can be understood as a decreasingly marginal willingness to pay for ESG.[38] For Republican-leaning areas, the estimator's bias due to sample selection is notable. Republican investors with ESG funds in their menu show high ESG fund inflow due to the scandals mainly because of the existence of ESG options. With all plans in the sample, it is the existence that matters rather than the scandal. This provides clear evidence that ESG scandals mainly have attention effects. The change in salience would easily cause a change in the portfolio composition but hardly a change in the weights if there is already an ESG fund in the portfolio.

1.6 Heterogeneous Leanings Towards Specific Elements: Decomposition by Investors

In this section, the attention shifts to investors leaning towards specific elements when choosing ESG funds to invest in.[39] We conduct a quantitative analysis of the learnings that align with heterogeneous investor preferences. Through this analysis, we also illustrate the mismatch between composite ESG scores and heterogeneous investor preferences. Although investors tend to decompose the indices, the goal is to emphasize the necessity for a broader array of ESG funds.

1.6.1 ESG Score Preferences in Fund Additions

As mentioned in Section 2, we acquire the environmental, social, and governance scores of each ESG fund added to the 401(k) plans. As illustrated in Table 1.8, we conduct a regression model as follows:

$$Y_{i,t} = \boldsymbol{\beta X_{i,t-1}} + \boldsymbol{\Gamma Controls_{i,t}} + \textbf{Fixed Effects}.$$

[38] As discussed by Heeb et al. (2023), plans featuring ESG options signify a primary commitment to ESG principles.

[39] Because of fiduciary concerns, most ESG funds in 401(k) plans are white-label composite ESG funds.

We examine the separate environmental (E), social (S), and governance (G) scores denoted as $Y_{i,t}$ of value-weighted ESG funds added to plan i in year t. If investors have specific preferences in terms of E/S/G scores, we will observe different βs of specific scores, introduced by different E/S/G scandals. ESG funds, centered around a common concept, show differences in asset holdings and benchmarks. Despite similar names, their investment strategies may vary. We utilize the separate E/S/G scores provided by MSCI for two primary reasons. Firstly, MSCI has maintained a relatively consistent ESG evaluation process since its reformation in 2012. Secondly, adhering to a single ESG evaluation system enables us to estimate investor preferences quantitatively.

In Table 1.15, we employ the environmental (E), social (S), and governance (G) scores from value-weighted asset holdings, normalized with the average score of each mutual fund category, as dependent variables. The results for the entire sample indicate that a 100-social-scandal increase corresponds to an approximately 0.2 increase in all E and S scores across ESG funds added to plans. The results are similar in Democratic, young and rich investors. This is mainly because they care about both social and environmental scandals and some scandals are attributed to these two categories simultaneously. Their preferences lead them to invest in funds with high scores of E and S. For the other group of investors, social scandals would lead to high social score funds being added to the portfolio. Compared with the sample mean, this magnitude would be around 33% to 50%. The cross effect is quite weak. Social scandals are less likely to cause them to choose high E-score funds.

The environmental scandals would also make investors add ESG funds with high E-scores. This only works for the group of investors who adjust portfolios due to environmental scandals, including Democratic and young investors. The magnitude of such leanings is a bit lower than the social aspect. One s.d. more environmental scandals would lead them to choose ESG funds with E-score 20-30% higher than average. Given no one cares about governance issues in general, we find no scandal would cause investors to invest in high G-score funds. In general, after governance scandals, investors prefer low E-score funds in some cases. This is possibly due to the fact that governance is a firm's internal agency problem and governance failure sometimes indicates pessimism in future externalities like carbon emissions in the whole industry.

The results in Table 1.15 reveal clear leanings towards certain elements in ESG based on investor preferences. This also raises questions about the adequacy of a conventional 1:1:1 ratio, which only partially accommodates the demands of ESG investing. Regarding the environmental component, the findings suggest that investors who care about this aspect allocate 30% more weight to it than to the social and governance components. Similarly, for the social dimension, investors respond by allocating up to 50% more than they do to other components. Recall that, in Section 4, older and wealthier investors react to the social dimension, whereas younger investors are more concerned about the environmental aspect. On the other hand, ESG data providers such as MSCI and KLD provide assessments weighted evenly across the three components, close to a 1:1:1 ratio. For example, KLD maintains a fixed ratio of 30.5:35.5:30 across E/S/G factors. However, despite the availability of separate E/S/G scores, decomposing such integrated indices when selecting funds to add is resource-intensive and time-consuming.

It is crucial to acknowledge that investors are indeed recognizing ESG funds with specific tilts. However, these findings only provide a conservative approximation of their actual preferences. These magnitudes are largely constrained by the availability of ESG funds with score weights closely resembling the 1:1:1 ratio. Given the relative scarcity of funds exclusively focusing on the environmental (E), social (S), or governance (G) components, investors' choices are limited to the existing ESG funds accessible within the market, particularly within the context of 401(k) plans. Hence, the magnitudes observed in this study represent the lower boundary of investors' inclinations.

1.6.2 ESG Scores on ESG Fund Flows

We extend the investigation to the existing fund flows, comparing the roles of the basic three ESG aspects. In particular, we aim to determine investor preferences for specific E/S/G scores in response to scandals correlated with fund flows. the regression model is structured as follows:

$$\begin{aligned}
Flow_Employee_{i,j,t} \quad = \quad & \beta_1 E_Score_{j,t} \times ESG_Scandal_{i,t-1} + \\
& \beta_2 S_Score_{j,t} \times ESG_Scandal_{i,t-1} + \\
& \beta_3 G_Score_{j,t} \times ESG_Scandal_{i,t-1} + \\
& \beta_0 ESG_Scandal'_{i,t-1} + \mathbf{\Gamma} \boldsymbol{Controls_{i,j,t}} + \textbf{Fixed Effects}.
\end{aligned}$$

Here, $E\_Score_{j,t}$, $S\_Score_{j,t}$, and $G\_Score_{j,t}$ denote the environmental, social, and governance scores, respectively, of fund j in year t. The variable $ESG\_Scandal'_{i,t-1}$ represents the logarithm of $ESG\_Scandal_{i,t-1}$ to mitigate multicollinearity. All other settings remain consistent with Table 1.11. The coefficients of the interaction terms capture the preference towards specific E/S/G scores triggered by ESG scandals.

The first column of Table 1.16 reveals specific preferences for ESG scores. It suggests a preference for ESG funds characterized by higher S-scores and lower G-scores, with a magnitude of nearly 1:-1. When considering political leanings, Republican investors tend to allocate more assets to high S-score and high E-score funds, with a difference of approximately 50%. As previously noted in Table 1.13, Republican-leaning investors displayed dual responsiveness to both environmental and social scandals, with a sensitivity ratio of 2:3. This highlights that investors align their decisions with their concerns. The negative coefficients of β_3 in Table 1.16 indicate that investors, including Democratic-leaning ones, tend to avoid investing in G-scores. The magnitude of the shift away from G-scores is approximately one-half of the preference for E-scores. This observation aligns with the finding that governance scandals have no impact on existing fund flows.

The age and wealth groups show similar results with political leanings. Only young and rich investors react to the scandals with a clear preference for ESG funds with high S-scores and low G-scores. The E-score has no significant effect on fund flows. The degrees of preference and aversion toward these two scores maintain a close 1:1 ratio. These coefficients of interaction signify the supplementary effects brought about by scandals. Generally, when ESG funds follow conventional ESG indices as benchmarks, investors tend to emphasize the social aspect while minimizing the governance aspect. The heterogeneous tilts to each score by investors match their preferences.

Investors are re-evaluating the composition of ESG funds based on the shocks they react to. On one hand, this provides evidence of the existing inconsistency between composite products and heterogeneous preferences. On the other hand, it challenges the need for more issue-specific ESG products. To refute this, the explanation can be divided into two parts. Firstly, the nature of composite indices may never fully align with the diverse demands of investors. Secondly, our discussion primarily focuses on the weighting aspect of assessments, potentially overlooking the measurement of each ESG factor. The measures of individual issues might not consistently reflect investor preferences either. While our analysis could not address this aspect, it remains a valid concern that cannot be disregarded.

1.7 Conclusion and Directions for Future Research

To understand how investors value ESG granularly, we document the heterogeneous investor preferences within their 401(k) investments in response to ESG scandals. We delve into three key dimensions of investor characteristics: political leaning, age, and wealth, revealing the heterogeneity in investment decisions in terms of new ESG fund additions and existing fund flows. In response to past nearby ESG scandals, plans in Democratic-leaning areas, predominantly populated by older and wealthier investors, tend to add ESG funds to their 401(k) menus. Upon categorizing these scandals, we find that social incidents, including human rights and cybersecurity issues, trigger fund additions among these investors. Younger investors also respond to environmental scandals by adding ESG funds, conditional on plan menu changes. These responses are primarily driven by political beliefs, followed by financial constraints and exposure to relevant scandals.

To examine existing ESG fund flows, we reveal a different perspective on heterogeneity. ESG scandals stimulate ESG fund inflows, primarily among Republican-leaning, young, and affluent investors. In terms of the triggers, both social and environmental scandals prompt ESG fund inflows, with economic and consumer rights concerns having the most significant impact.

To illustrate investors leaning towards specific elements in ESG, we quantified their preferences across the environmental, social, and governance components of ESG investments. In the ESG fund-addition case, investors display a preference for those with high S- and E-scores, which are up to 50% and 30% more preferred than those with high G-scores, respectively. This preference remains consistent when considering fund flows, with the inclination towards S-score funds even doubling in certain subsample groups. Investors exhibit varying preferences for each score in terms of magnitude.

These findings provide a feature-based demand for specific ESG funds. Although numerous investor characteristics contribute to heterogeneity, the three we studied could be combined in some dimensions, providing industrial insights for fund marketing. In addition, we provide an empirical foundation for the theoretical study of the trade-off between cognitive cost and non-pecuniary utility, leading to the study of investor welfare improvement. As shown in this study, people have different leanings towards each element, and they invest in composite ESG products in their 401(k) plans. They are saving cognitive costs and bear some non-pecuniary utility loss due to the mismatch. We acknowledge that this study should be combined with future quantitative models to unravel the endogeneity of ESG demands. These models could quantify investor utility losses and offer guidance for social planners. Ultimately, our objective is to emphasize the heterogeneous role of each ESG element, advocating for a more comprehensive approach that accounts for diverse preferences.

Bibliography

Agnew, Julie, Pierluigi Balduzzi, and Annika Sunden. 2003. "Portfolio Choice and Trading in a Large 401 (k) Plan." *American Economic Review*, 93(1): 193–215.

Allcott, Hunt, Matthew Gentzkow, and Lena Song. 2022. "Digital Addiction." *American Economic Review*, 112(7): 2424–2463.

Alok, Shashwat, Nitin Kumar, and Russ Wermers. 2020. "Do Fund Managers Misestimate Climatic Disaster Risk." *The Review of Financial Studies*, 33(3): 1146–1183.

Baker, Malcolm, Mark L Egan, and Suproteem K Sarkar. 2022. "How Do Investors Value ESG?" National Bureau of Economic Research Working Paper 30708.

Baldauf, Markus, Lorenzo Garlappi, and Constantine Yannelis. 2020. "Does Climate Change Affect Real Estate Prices? Only If You Believe in It." *The Review of Financial Studies*, 33(3): 1256–1295.

Barber, Brad M, Adair Morse, and Ayako Yasuda. 2021. "Impact Investing." *Journal of Financial Economics*, 139(1): 162–185.

Bauer, Rob, Tobias Ruof, and Paul Smeets. 2021. "Get Real! Individuals Prefer More Sustainable Investments." *The Review of Financial Studies*, 34(8): 3976–4043.

Benartzi, Shlomo, and Richard H Thaler. 2001. "Naive Diversification Strategies in Defined Contribution Saving Plans." *American Economic Review*, 91(1): 79–98.

Berk, Jonathan, and Jules H Van Binsbergen. 2021. "The impact of impact investing." Stanford University Graduate School of Business and Law & Economics Center at George Mason University Scalia Law School Research Paper.

Bernstein, Asaf, Matthew T Gustafson, and Ryan Lewis. 2019. "Disaster on the Horizon: The Price Effect of Sea Level Rise." *Journal of Financial Economics*, 134(2): 253–272.

Bolton, Patrick, and Marcin Kacperczyk. 2021. “Do Investors Care About Carbon Risk?” *Journal of Financial Economics*, 142(2): 517–549.

Bradbury, Meike AS, Thorsten Hens, and Stefan Zeisberger. 2015. “Improving Investment Decisions with Simulated Experience.” *Review of Finance*, 19(3): 1019–1052.

Broccardo, Eleonora, Oliver Hart, and Luigi Zingales. 2022. “Exit versus Voice.” *Journal of Political Economy*, 130(12): 3101–3145.

Chapman, David A, and Valery Polkovnichenko. 2009. “First-Order Risk Aversion, Heterogeneity, and Asset Market Outcomes.” *The Journal of Finance*, 64(4): 1863–1887.

Chava, Sudheer. 2014. “Environmental Externalities and Cost of Capital.” *Management Science*, 60(9): 2223–2247.

Choi, James J. 2015. “Contributions to Defined Contribution Pension Plans.” *Annual Review of Financial Economics*, 7: 161–178.

Choi, James J, David Laibson, and Andrew Metrick. 2002. “How Does the Internet Affect Trading? Evidence from Investor Behavior in 401 (k) Plans.” *Journal of Financial Economics*, 64(3): 397–421.

Christensen, Dane M, George Serafeim, and Anywhere Sikochi. 2022. “Why Is Corporate Virtue in the Eye of the Beholder? The Case of ESG Ratings.” *The Accounting Review*, 97(1): 147–175.

Cohen, Alma, and Liran Einav. 2007. “Estimating Risk Preferences from Deductible Choice.” *American Economic Review*, 97(3): 745–788.

Cohen, Lauren, and Breno Schmidt. 2009. “Attracting Flows by Attracting Big Clients.” *The Journal of Finance*, 64(5): 2125–2151.

Cohen, Lauren, Umit G Gurun, and Quoc H Nguyen. 2020. “The ESG-Innovation Disconnect: Evidence from Green Patenting.” National Bureau of Economic Research Working Paper 27990.

Davis, Gerald F, and E Han Kim. 2007. “Business Ties and Proxy Voting by Mutual Funds.” *Journal of Financial Economics*, 85(2): 552–570.

Döttling, Robin, and Sehoon Kim. 2022. "Sustainability Preferences Under Stress: Evidence from COVID-19." *Journal of Financial and Quantitative Analysis*, 1–39.

Engle, Robert F, Stefano Giglio, Bryan Kelly, Heebum Lee, and Johannes Stroebel. 2020. "Hedging Climate Change News." *The Review of Financial Studies*, 33(3): 1184–1216.

Goldstein, Itay, Alexandr Kopytov, Lin Shen, and Haotian Xiang. 2022. "On ESG Investing: Heterogeneous Preferences, Information, and Asset Prices." National Bureau of Economic Research Working Paper 29839.

Gompers, Paul, Anna Kovner, Josh Lerner, and David Scharfstein. 2005. "Venture Capital Investment Cycles: The Role of Experience and Specialization." *Journal of Financial Economics*, 81(1): 649–679.

Hartzmark, Samuel M, and Abigail B Sussman. 2019. "Do Investors Value Sustainability? A Natural Experiment Examining Ranking and Fund Flows." *The Journal of Finance*, 74(6): 2789–2837.

Heckman, James J. 1979. "Sample Selection Bias as a Specification Error." *Econometrica*, 47(1): 153–161.

Heeb, Florian, Julian F Kölbel, Falko Paetzold, and Stefan Zeisberger. 2023. "Do Investors Care about Impact?" *The Review of Financial Studies*, 36(5): 1737–1787.

Heinkel, Robert, Alan Kraus, and Josef Zechner. 2001. "The Effect of Green Investment on Corporate Behavior." *Journal of Financial and Quantitative Analysis*, 36(4): 431–449.

Huberman, Gur, and Wei Jiang. 2006. "Offering versus Choice in 401 (k) Plans: Equity Exposure and Number of Funds." *The Journal of Finance*, 61(2): 763–801.

Humphrey, Jacquelyn, Shimon Kogan, Jacob Sagi, and Laura Starks. 2021. "The Asymmetry in Responsible Investing Preferences." National Bureau of Economic Research Working Paper 29288.

Kashyap, Anil K, Natalia Kovrijnykh, Jian Li, and Anna Pavlova. 2023. "Is there too much benchmarking in asset management?" *American Economic Review*, 113(4): 1112–1141.

Liao, Li, Zhengwei Wang, Jia Xiang, Hongjun Yan, and Jun Yang. 2021. "User Interface and Firsthand Experience in Retail Investing." *The Review of Financial Studies*, 34(9): 4486–4523.

Li, Shangchen, Hongxun Ruan, Sheridan Titman, and Haotian Xiang. 2023. "ESG Spillovers." National Bureau of Economic Research Working Paper 31248.

Ma, Liye, Ramayya Krishnan, and Alan L Montgomery. 2015. "Latent homophily or social influence? An empirical analysis of purchase within a social network." *Management Science*, 61(2): 454–473.

Murfin, Justin, and Matthew Spiegel. 2020. "Is the Risk of Sea Level Rise Capitalized in Residential Real Estate?" *The Review of Financial Studies*, 33(3): 1217–1255.

Pástor, Luboš, Robert F Stambaugh, and Lucian A Taylor. 2021. "Sustainable Investing in Equilibrium." *Journal of Financial Economics*, 142(2): 550–571.

Pavlova, Anna, and Taisiya Sikorskaya. 2023. "Benchmarking intensity." *The Review of Financial Studies*, 36(3): 859–903.

Pool, Veronika K, Clemens Sialm, and Irina Stefanescu. 2016. "It Pays to Set the Menu: Mutual Fund Investment Options in 401 (k) Plans." *The Journal of Finance*, 71(4): 1779–1812.

Pool, Veronika K, Clemens Sialm, and Irina Stefanescu. 2022. "Mutual Fund Revenue Sharing in 401 (k) Plans." National Bureau of Economic Research Working Paper 30721.

Renneboog, Luc, Jenke Ter Horst, and Chendi Zhang. 2011. "Is Ethical Money Financially Smart? Nonfinancial Attributes and Money Flows of Socially Responsible Investment Funds." *Journal of Financial Intermediation*, 20(4): 562–588.

Riedl, Arno, and Paul Smeets. 2017. "Why Do Investors Hold Socially Responsible Mutual Funds?" *The Journal of Finance*, 72(6): 2505–2550.

Sialm, Clemens, Laura Starks, and Hanjiang Zhang. 2015*a*. "Defined Contribution Pension Plans: Mutual Fund Asset Allocation Changes." *American Economic Review*, 105(5): 432–436.

Sialm, Clemens, Laura Starks, and Hanjiang Zhang. 2015*b*. "Defined Contribution Pension Plans: Sticky or Discerning Money?" *The Journal of Finance*, 70(2): 805–838.

Sialm, Clemens, Laura Starks, and Hanjiang Zhang. 2018. "Defined Contribution Pension Plans and Mutual Fund Flows." *Journal of Investment Management*, 16(3): 31–45.

Tang, Ning, Olivia S Mitchell, Gary R Mottola, and Stephen P Utkus. 2010. "The Efficiency of Sponsor and Participant Portfolio Choices in 401 (k) Plans." *Journal of Public Economics*, 94(11-12): 1073–1085.

Tarrant, Mark, Adrian C North, and David J Hargreaves. 2001. "Social categorization, self-esteem, and the estimated musical preferences of male adolescents." *The Journal of Social Psychology*, 141(5): 565–581.

Tran, Anh, and Pingle Wang. 2023. "Barking up the Wrong Tree: Return-Chasing in 401 (k) Plans." *Journal of Financial Economics*, 148(1): 69–90.

Wu, Yue, Kaifu Zhang, and Jinhong Xie. 2020. "Bad Greenwashing, Good Greenwashing: Corporate Social Responsibility and Information Transparency." *Management Science*, 66(7): 3095–3112.

Xiouros, Costas, and Fernando Zapatero. 2010. "The Representative Agent of an Economy with External Habit Formation and Heterogeneous Risk Aversion." *The Review of Financial Studies*, 23(8): 3017–3047.

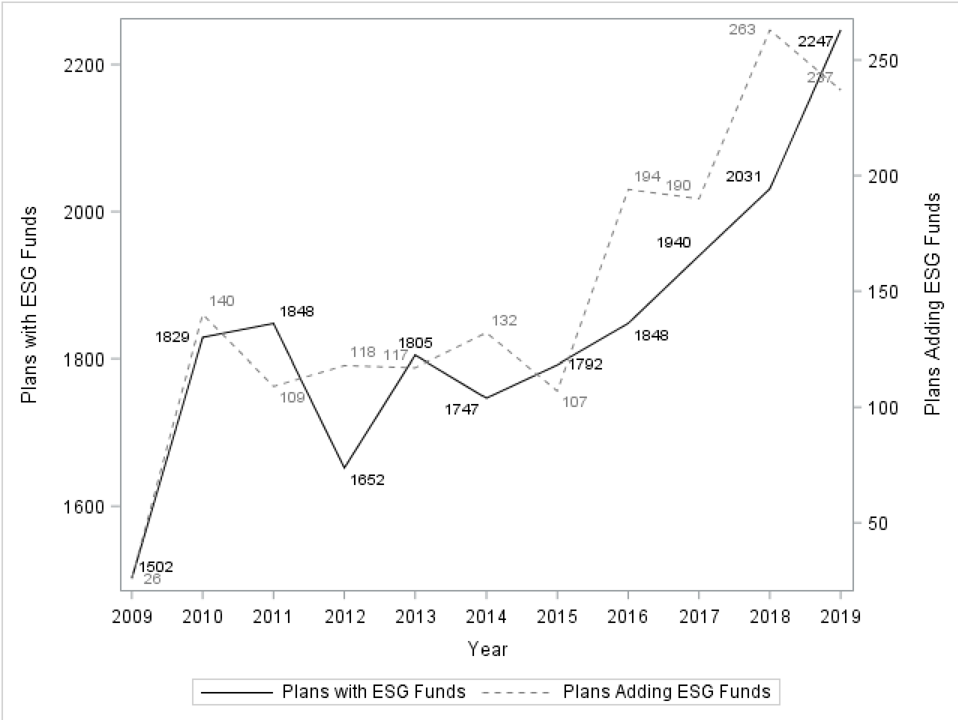
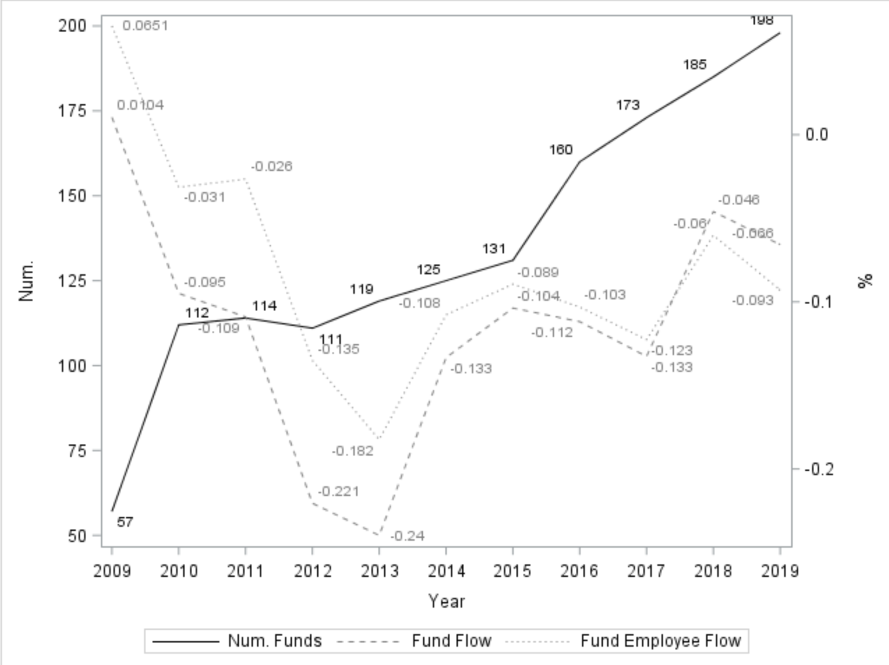
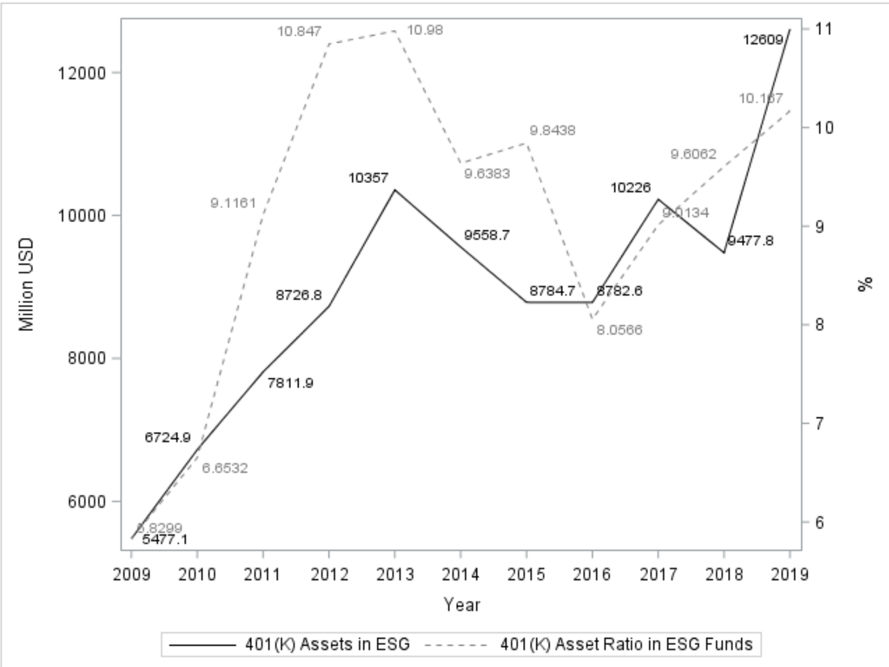


Figure 1.1: ESG investment in 401(k) plans



(a) ESG funds flow in 401(k) plans



(b) ESG investment volume and assets held in 401(k) plans

Figure 1.2: ESG funds in 401(k) plans

Table 1.1: Summary statistics

The table presents summary statistics for the whole sample observations. Panels (a) and (b) display observations at the plan-year level, while Panel (c) shows fund-plan-year observations. The sample includes all plans with total assets greater than 10M USD and that have been set up for more than five years. Merged and closed plans are also included to avoid survival bias. Panel (a) provides summary statistics for variables related to fund characteristics, while Panel (b) presents scandal-related variables. Panel (c) focuses on all ESG funds within our sample plans.

(a) Plans

	N	Mean	Std. Dev.	p5	Median	p95
ESG_Funds	124827	0.121	0.416	0	0	1
ESG_Add	117611	0.013	0.152	0	0	0
Size	117611	17.578	1.224	16.226	17.258	20.066
MenuSize	117611	3.271	0.351	2.708	3.296	3.761
AutoEnroll	117611	0.400	0.490	0	0	1
DefaultInv	117611	0.867	0.339	0	1	1
DualRole	117611	0.633	0.482	0	1	1
Geo_ESG	122525	4.585	0.949	2.918	4.595	6.190
R_Vote	124831	0.525	0.135	0.287	0.524	0.738
D_Vote	124831	0.408	0.130	0.212	0.405	0.627
$Flow_{TDF\leq 2040}/Flow_{TDF>2040}$	89719	1.22	5.092	-3.086	0.797	6.569
TDF_Age	98949	48.979	4.109	41.904	49.253	55.073
Account_Balance	119171	2.393	0.059	2.288	2.397	2.486
Deferral	111219	5213.886	2696.46	1787.506	4610.767	10800.219

Table 1.1: Summary statistics

(b) Scandals

	N	Mean	Std. Dev.	p5	Median	p95
ESG_Scandal	117611	1.066	1.383	0.02	0.55	4.08
Environment_Scandal	117611	0.308	0.468	0	0.13	1.32
Social_Scandal	117611	0.605	0.887	0.01	0.28	2.49
Governance_Scandal	117611	0.648	1.560	0	0.2	2.63
S_Environment	117611	0.096	0.152	0	0.04	0.45
S_Social	117611	0.330	0.547	0	0.14	1.53
S_Governance	117611	0.506	1.367	0	0.14	1.78
Animal & Biodiversity	117611	1.843	3.503	0	0	9
Cyber & Privacy Concerns	117611	7.028	19.106	0	1	46
Economic & Consumer Rights	117611	5.047	9.641	0	1	26
Environmental Issues	117611	19.224	31.777	0	7	94
Health & Safety	117611	8.955	14.039	0	4	38
Labor & Employment	117611	6.308	10.713	0	2	29
Legal & Compliance	117611	4.812	8.779	0	1	21
Political & Governance	117611	0.578	1.574	0	0	3
Resource & Infrastructure	117611	2.366	3.751	0	1	10
Social & Human Rights	117611	7.563	15.022	0	2	37
Substance & Social Ethics	117611	1.448	4.161	0	0	9
Technology & Surveillance	117611	0.060	0.278	0	0	0
Weapons & Security	117611	1.745	4.308	0	0	7

(c) ESG funds

	N	Mean	Std. Dev.	p5	Median	p95
Flow_Employee	46383	-0.114	0.561	-1.66	-0.001	0.745
$Ret_{t-1,t-3}$	46383	0.243	0.252	-0.203	0.272	0.606
$Vol_{t-1,t-3}$	46383	1.640	0.655	0.742	1.509	2.766

Table 1.2: ESG funds per plan by different features

The table presents the average number of ESG funds per plan within different characteristic groups. The first grouping variable is the ESG score of nearby public firms within the same two-digit zip code area. The political leaning groups are defined using the presidential voting results. Ages are based on the measure $Flow_{TDF\leq 2040}/Flow_{TDF>2040}$. Wealth is represented by the annual deferral, which is the participant's contribution. Each characteristic, except the political leaning, is divided into two groups based on the median value and annually rebalanced. The differences and t-statistics are reported in the last two columns.

Features	ESG Funds per Plan			
	Low	High	Diff	t
Geo_ESG	14.93%	14.52%	0.41	2.21
	Republican	Democratic	Diff	t
Political Leaning	13.73%	17.81%	-4.07	-20.97
	Old	Young	Diff	t
Age	13.76%	15.97%	-2.21	-12.02
	Poor	Rich	Diff	t
Wealth	13.34%	15.52%	-2.18	-11.96

Table 1.3: Plan menu changes triggered by ESG scandals

The table presents the estimated coefficients for the plan-year-level panel model $Y_{i,t} = \beta ESG\_Scandal_{i,t-1} + \mathbf{\Gamma} \boldsymbol{Controls_{i,t}} + \textbf{Fixed Effects}$, where $ESG\_Scandal_{i,t-1}$ represents the ESG scandals in the first-two-digit zip code area in the preceding year $(t-1)$ of fund i. $Y_{i,t}$ denotes the menu changes according to the ESG funds at the year t, where ESG_Fund represents the percentage of ESG funds in each plan, ESG_Add and ESG_Del are the ESG funds added and deleted in such plan normalized to the menu size this year. ESG_In and ESG_Out are the ESG funds replaced into and out of such plan normalized to the menu size this year. The variable *Controls* includes the following covariates: log plan size ($Size$), log menu size ($MenuSize$), an indicator for employee auto-enrollment ($AutoEnroll$), an indicator for TDF as the default investment option ($DefaultInv$), and an indicator for the mutual fund company serving as the plan's main service provider ($DuralRole$). The table includes year indicators as control variables as well. The standard errors are clustered at the plan and year level to account for potential correlations. Statistical significance levels are denoted by *, **, and ***, corresponding to the 10%, 5%, and 1% significance levels, respectively.

	(1) ESG_Fund	(2) ESG_Add	(3) ESG_Del	(4) ESG_In	(5) ESG_Out	(6) ESG_Add	(7) ESG_Add	(8) ESG_Add
ESG_Scandal$_{i,t-1}$	**0.005****	**0.014****	0.002	-0.000	0.000	**0.224*****	**0.254*****	**0.524****
	(0.002)	**(0.007)**	(0.005)	(0.002)	(0.002)	**(0.078)**	**(0.073)**	**(0.222)**
Observations	114,645	4,316	7,601	27,468	27,468	7,222	114,645	5,595
R-squared	0.757	0.488	0.460	0.435	0.411			
Reg Method	OLS	OLS	OLS	OLS	OLS	Logit	Logit	Logit
Conditional	Uncon	Adding	Deleting	Replacing	Replacing	Change	Uncon	Change
Cross-Sectional Fixed Effect	Plan	Plan	Plan	Plan	Plan	Plan	Geo	Geo

Table 1.4: Mechanism of scandal impacts on portfolio changes

The following table presents the estimated coefficients for the plan-year level panel model similar in Table 1.3. The *ESG_Scandals* is substituted by *HS* as the number of high-severity scandals, *HR* as the number of high-reach scandals, or *New* as the number of new scandals. The definitions of severity, reach and new are all from RepRisk. The conditional cases are indicated at the column heads. All regressions include plan and year fixed effects. All standard errors are clustered at the plan and year level to account for potential correlations. Levels of statistical significance are denoted by *, **, and ***, corresponding to the 10%, 5%, and 1% significance levels, respectively.

Table 1.5: Severity vs. Reach vs. Novelty

	(1)	(2)	(3)	(4)	(5)	(6)
Condition		Unconditional			Change	
HS_{t-1}	-0.196			0.028		
	(0.185)			(0.046)		
HR_{t-1}		0.040**			0.005*	
		(0.020)			(0.003)	
New_{t-1}			0.043*			0.010**
			(0.023)			(0.004)

Table 1.6: Heterogeneity in fund additions

The following three panels present the estimated coefficients for the plan-year-level panel model in Table 1.3 and the logit model Prob($ESG\_Add_{i,t} > 0$) $= \Lambda(\beta ESG\_Scandal_{i,t-1} + \mathbf{\Gamma} \boldsymbol{Controls_{i,t}} + \textbf{Fixed Effects})$. $ESG\_Scandal_{i,t-1}$ represents the occurrence of ESG scandals in the first-two-digit zip code area of plan i during the preceding year $(t-1)$. The variable $Controls_{i,t}$ remains consistent with Table 1.3. Additionally, year indicators are included as control variables in the analysis. In Panel (a), the sample is divided into the Democratic and Republican areas. Moving to Panel (b), the average employee age is estimated using two variables: $Flow_{TDF\leq 2040}/Flow_{TDF>2040}$ and the TDF Age. In Panel (c), employee wealth is proxied by the average employee contribution (Deferral) and the average account balance (Account Balance). Each pair of subsamples is rebalanced annually. Median values in each year are used in Panels (b) and (c) to have the subsamples of each column. Standard errors are clustered at the plan and year level to account for potential correlations. Levels of statistical significance are denoted by *, **, and ***, corresponding to the 10%, 5%, and 1% significance levels, respectively.

(a) Political Voting

VARIABLES	(1) Democratic	(2) Republican	(3) Democratic	(4) Republican	(5) Democratic	(6) Republican	(7) Democratic	(8) Republican
$ESG\_Scandal_{i,t-1}$	**0.281*** **	0.116	**0.250** **	0.199	**0.029** **	-0.001	**0.002** **	-0.000
	(0.095)	(0.129)	**(0.124)**	(0.148)	**(0.015)**	(0.005)	**(0.001)**	(0.002)
Observations	35,439	78,872	1,931	3,596	1,191	3,097	33,937	77,279
Reg Method	Logit	Logit	Logit	Logit	OLS	OLS	OLS	OLS
Conditional	Uncon	Uncon	Change	Change	Uncon	Uncon	Change	Change
Cross-Section Fixed Effect	Geo	Geo	Geo	Geo	Plan	Plan	Plan	Plan

(b) Age

	Flow Age				TDF Age			
VARIABLES	Young	Old	Young	Old	Young	Old	Young	Old
$ESG\_Scandal_{i,t-1}$	**0.172***	**0.376*** **	0.115	**0.209** **	**0.188** **	**0.334*** **	0.159	0.198
	(0.104)	**(0.111)**	(0.148)	**(0.092)**	**(0.095)**	**(0.118)**	(0.112)	(0.153)
Observations	56,399	56,236	1,985	2,045	52,547	60,109	2,262	2,606
Reg Method	Logit	Logit	Logit	Logit	Logit	Logit	Logit	Logit
Conditional	Uncon	Uncon	Change	Change	Uncon	Uncon	Change	Change
Cross-Section Fixed Effect	Geo	Geo	Geo	Geo	Geo	Geo	Geo	Geo

Table 1.6: Heterogeneity in fund additions

(c) Wealth

	Deferral				Account Balance			
VARIABLES	Poor	Rich	Poor	Rich	Poor	Rich	Poor	Rich
$\text{ESG\_Scandal}_{i,t-1}$	**0.380***	**0.202****	**0.443***	**0.269*****	0.127	**0.256*****	0.327	**0.175***
	(0.197)	**(0.086)**	**(0.263)**	**(0.103)**	(0.146)	**(0.090)**	(0.231)	**(0.090)**
Observations	34,828	63,190	1,295	2,785	28,556	77,878	1,235	3,298
Reg Method	Logit	Logit	Logit	Logit	Logit	Logit	Logit	Logit
Conditional	Uncon	Uncon	Change	Change	Uncon	Uncon	Change	Change
Cross-Section Fixed Effect	Geo	Geo	Geo	Geo	Geo	Geo	Geo	Geo

Table 1.7: The priority role of political leanings in ESG fund additions

The table presents the estimated coefficients βs for the plan-year-level panel logit model $\text{Prob}(ESG\_Add_{i,t} > 0) = \Lambda(\beta ESG\_Scandal_{i,t-1} + \mathbf{\Gamma}\boldsymbol{Controls_{i,t}} + \textbf{Fixed Effects})$. $ESG\_Scandal_{i,t-1}$ represents the incidence of ESG scandals in the first-two-digit area of plan i during the preceding year $(t-1)$. $ESG\_Add_{i,t}$ denotes the proportion of ESG funds added to each plan, normalized to the menu size for the current year. The variable $Controls_{i,t}$ remains consistent with the values in Table 1.3. Additionally, the table includes year and county indicators as control variables. Each pair of coefficients and standard errors corresponds to a single regression conducted within specific subsamples, determined by a double-sorting mechanism based on presidential voting results and either wealth or age. The proxies for wealth and age follow the definitions outlined in Table 1.6. In the upper panel, all plans, including those without changing their menus, are included in the regression, while the lower panel focuses only on plans that have changed menus during year t. To account for potential within-plan correlation, standard errors are clustered at the plan level. Statistical significance levels are denoted by *, **, and ***, corresponding to the 10%, 5%, and 1% significance levels, respectively.

Condition			Republican		Democratic	
			Coeff.	S.E.	Coeff.	S.E.
Unconditional	Deferral	Poor	0.332	(1.366)	**0.430***	**(0.259)**
		Rich	-0.018	(0.106)	**0.205***	**(0.111)**
	Account Balance	Poor	0.075	(0.311)	0.042	(0.191)
		Rich	0.118	(0.733)	**0.315*****	**(0.118)**
	Flow Age	Young	0.163	(0.689)	0.213	(0.156)
		Old	0.065	(0.297)	**0.246***	**(0.149)**
	TDF Age	Young	0.241	(1.243)	0.154	(0.135)
		Old	0.051	(0.208)	**0.352****	**(0.168)**
Change	Deferral	Poor	-0.281	(0.708)	-0.172	(0.728)
		Rich	0.973	(0.663)	**0.777****	**(0.309)**
	Account Balance	Poor	-1.971*	(1.092)	0.265	(0.643)
		Rich	**1.141***	**(0.677)**	**0.716****	**(0.324)**
	Flow Age	Young	-0.171	(0.797)	0.225	(0.425)
		Old	-1.140	(1.180)	**1.175*****	**(0.450)**
	TDF Age	Young	-1.437	(1.072)	0.369	(0.363)
		Old	-0.147	(1.003)	0.165	(0.579)

Table 1.8: Responses to environmental, social, or governance scandals

The table presents the estimated coefficients βs for the plan-year-level panel logit model $\text{Prob}(ESG\_Add_{i,t} > 0) = \Lambda(\beta X_{i,t} + \mathbf{\Gamma}\,\boldsymbol{Controls_{i,t}} + \textbf{Fixed Effects})$, where $ESG\_Add_{i,t}$ denotes the proportion of ESG funds added to plan i, normalized to the menu size for the current year. The independent variable $X_{i,t}$ comprises the number of social scandals ($Social_{i,t-1}$), environmental scandals ($Environ_{i,t-1}$), and governance scandals ($Gover_{i,t-1}$). Each column represents one separate regression. To address multicollinearity concerns, scandals attributed exclusively to the social aspect ($S\_Social_{i,t-1}$), environmental aspect ($S\_Environ_{i,t-1}$), and governance aspect ($S\_Gover_{i,t-1}$) are included jointly in the regression model. The variable $Controls_{i,t}$ remains consistent with the values in Table 1.3. Additionally, the table incorporates year and county indicators as control variables to account for potential temporal and regional effects. To address potential correlations, standard errors are clustered at the plan and year level. Statistical significance levels are denoted by *, **, and ***, corresponding to the 10%, 5%, and 1% significance levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Condition		Unconditional				Change		
$\text{Social}_{i,t-1}$	**0.368*****				**0.821*****			
	(0.088)				**(0.276)**			
$\text{Environ}_{i,t-1}$		-0.013				0.739		
		(0.230)				(0.629)		
$\text{Gover}_{i,t-1}$			0.052				0.048	
			(0.063)				(0.161)	
$\text{S\_Social}_{i,t-1}$				**0.486*****				**1.115*****
				(0.116)				**(0.351)**
$\text{S\_Environ}_{i,t-1}$				-0.396				2.877
				(0.680)				(2.026)
$\text{S\_Gover}_{i,t-1}$				-0.013				-0.076
				(0.068)				(0.176)
Observations	114,645	114,645	114,645	114,645	7,222	7,222	7,222	7,222

Table 1.9: Heterogeneity in environmental, social, and governance scandals

The table presents the estimated coefficients βs for the plan-year-level panel logit model Prob$(ESG_Add_{i,t} > 0) = \Lambda(\beta X_{i,t} + \mathbf{\Gamma}\,\boldsymbol{Controls_{i,t}} + \textbf{Fixed Effects})$, where $ESG\_Add_{i,t}$ denotes the proportion of ESG funds added to plan i, normalized to the menu size for the current year. The independent variable $X_{i,t}$ is shown in each column header. The variable $Controls_{i,t}$ remains consistent with the values in Table 1.3. Each pair of coefficients and standard errors corresponds to a single regression conducted within specific subsamples shown in the first two columns. The subsample determinations are identical to those presented in Table 1.6. Additionally, the table includes year and county indicators as control variables. In the upper panel, all plans, including those without changing their menus, are included in the regression, while the lower panel focuses only on plans that have changed menus during that year. To account for potential within-plan correlation, standard errors are clustered at the plan level. Statistical significance levels are denoted by *, **, and ***, corresponding to the 10%, 5%, and 1% significance levels, respectively.

Unconditional							
		Social_{t-1}		Environ_{t-1}		Gover_{t-1}	
Political	Republican	0.109	(0.210)	0.014	(0.307)	0.116	(0.272)
	Democratic	**0.345*** **	**(0.111)**	-0.300	(0.347)	0.088	(0.072)
Flow Age	Old	**0.537*** **	**(0.136)**	-0.055	(0.365)	**0.326*** **	**(0.112)**
	Young	**0.230***	**(0.125)**	-0.145	(0.329)	-0.025	(0.077)
TDF Age	Old	**0.507*** **	**(0.146)**	0.085	(0.378)	0.208*	(0.107)
	Young	**0.257** **	**(0.114)**	-0.087	(0.289)	-0.020	(0.079)
Deferral	Poor	**0.521** **	**(0.250)**	-0.293	(0.490)	-0.058	(0.134)
	Rich	**0.309*** **	**(0.099)**	0.064	(0.289)	0.077	(0.078)
Account Balance	Poor	**0.321***	**(0.179)**	0.406	(0.401)	-0.134	(0.121)
	Rich	**0.332*** **	**(0.104)**	-0.255	(0.292)	0.125	(0.081)
Change							
		Social_{t-1}		Environ_{t-1}		Gover_{t-1}	
Political	Republican	-0.178	(0.929)	-0.948	(0.718)	-1.367	(0.963)
	Democratic	**1.016*** **	**(0.371)**	**1.589***	**(0.861)**	0.020	(0.181)
Flow Age	Old	**1.322** **	**(0.640)**	-0.434	(0.702)	**1.135** **	**(0.483)**
	Young	**0.565***	**(0.320)**	**1.038***	**(0.574)**	-0.020	(0.172)
TDF Age	Old	0.717	(0.603)	1.645	(1.468)	0.155	(0.343)
	Young	**0.838** **	**(0.336)**	**0.566** **	**(0.221)**	0.022	(0.182)
Deferral	Poor	-0.447	(1.668)	-2.531	(3.732)	-0.623	(1.617)
	Rich	**0.823*** **	**(0.317)**	0.804	(0.665)	0.159	(0.163)
Account Balance	Poor	0.603	(0.654)	**2.978** **	**(1.414)**	-0.344	(0.425)
	Rich	**0.954*** **	**(0.288)**	0.314	(0.698)	0.186	(0.181)

Table 1.10: Fund addition heterogeneity in responses to specific scandal issues

The following two panels present the estimated coefficients for the plan-year-level panel model in Table 1.3 and the logit model $\text{Prob}(ESG\_Add_{i,t} > 0) = \Lambda(\beta X_{i,t-1} + \boldsymbol{\Gamma} \boldsymbol{Controls_{i,t}} + \textbf{Fixed Effects})$, where $X_{i,t-1}$ represents the occurrence of specific ESG scandals near plan i, indicated in each column heading, during the preceding year $(t-1)$. The ESG scandals are categorized into 13 categories from RepRisk's dataset. The regression settings and subsample determinations are consistent with those in Table 1.9. Panel (a) includes all plans, even those that did not change their menus, in the regression analysis. In contrast, Panel (b) focuses solely on plans that modified their menus during the given year. To address potential correlations, standard errors are clustered at the plan and year level. Statistical significance levels are denoted by *, **, and ***, corresponding to significance at the 10%, 5%, and 1% levels, respectively.

(a) Unconditional case: All plan observations

		Animal & Biodiversity	Cyber & Privacy Concerns	Economic & Consumer Rights	Environmental Issues	Health & Safety	Labor & Employment	Legal & Compliance	Political & Governance	Resource & Infrastructure	Social & Human Rights	Substance & Social Ethics	Technology & Surveillance	Weapons & Security
Whole Sample		0.015 (0.018)	0.008*** (0.002)	-0.004 (0.009)	-0.001 (0.003)	-0.007 (0.006)	0.010* (0.005)	0.020*** (0.005)	0.052*** (0.018)	-0.013 (0.013)	0.013*** (0.003)	0.006 (0.011)	0.137 (0.100)	0.016 (0.011)
Political	Republican	-0.013 (0.025)	0.005 (0.005)	0.022 (0.015)	-0.000 (0.004)	0.002 (0.008)	0.009 (0.011)	-0.001 (0.011)	0.118** (0.055)	0.001 (0.017)	-0.000 (0.007)	-0.013 (0.016)	0.029 (0.171)	0.021 (0.031)
	Democratic	0.055** (0.025)	0.007*** (0.002)	-0.019* (0.011)	-0.003 (0.004)	-0.027** (0.011)	0.009 (0.007)	0.022*** (0.007)	0.031 (0.022)	-0.033 (0.021)	0.015*** (0.004)	0.011 (0.020)	0.161 (0.131)	0.003 (0.012)
Flow Age	Old	-0.017 (0.028)	0.009*** (0.003)	0.013 (0.015)	-0.003 (0.005)	-0.004 (0.009)	0.024*** (0.008)	0.017** (0.009)	0.023 (0.039)	-0.018 (0.019)	0.017*** (0.005)	0.014 (0.018)	0.283* (0.152)	-0.020 (0.022)
	Young	0.035 (0.024)	0.007** (0.003)	-0.015 (0.013)	-0.001 (0.004)	-0.010 (0.009)	-0.001 (0.008)	0.020*** (0.007)	0.052** (0.021)	-0.013 (0.019)	0.010** (0.004)	-0.001 (0.014)	0.021 (0.132)	0.025* (0.013)
TDF Age	Old	0.002 (0.026)	0.009** (0.004)	-0.007 (0.014)	0.001 (0.005)	-0.008 (0.010)	0.009 (0.008)	0.023*** (0.009)	0.038 (0.035)	-0.014 (0.019)	0.014*** (0.005)	0.009 (0.016)	0.139 (0.158)	-0.000 (0.019)
	Young	0.022 (0.025)	0.007** (0.003)	0.001 (0.012)	-0.002 (0.004)	-0.004 (0.008)	0.010 (0.007)	0.016** (0.007)	0.053** (0.022)	-0.009 (0.018)	0.011** (0.004)	0.007 (0.017)	0.127 (0.131)	0.022 (0.014)
Deferral	Poor	0.018 (0.044)	0.019*** (0.006)	-0.006 (0.022)	-0.006 (0.007)	-0.020 (0.015)	0.008 (0.011)	0.037*** (0.014)	0.100** (0.050)	-0.002 (0.031)	0.021** (0.009)	0.002 (0.023)	0.226 (0.235)	0.020 (0.024)
	Rich	-0.003 (0.021)	0.005** (0.002)	-0.004 (0.011)	0.001 (0.004)	0.001 (0.007)	0.009 (0.007)	0.016** (0.007)	0.056** (0.022)	-0.003 (0.016)	0.010*** (0.004)	0.011 (0.016)	0.097 (0.121)	0.012 (0.013)
Account Balance	Poor	0.039 (0.036)	0.006 (0.004)	0.017 (0.017)	0.005 (0.005)	-0.011 (0.014)	0.002 (0.010)	0.024** (0.010)	0.051 (0.032)	0.015 (0.026)	0.011* (0.006)	0.019 (0.020)	-0.139 (0.225)	0.046** (0.022)
	Rich	-0.001 (0.021)	0.008*** (0.002)	-0.008 (0.011)	-0.003 (0.004)	-0.004 (0.007)	0.013* (0.007)	0.017** (0.007)	0.054** (0.024)	-0.020 (0.016)	0.012*** (0.004)	0.004 (0.016)	0.234** (0.115)	0.001 (0.014)

Table 1.10: Fund addition heterogeneity in responses to specific scandal issues

(b) Conditional case: Plan observations with menu change

		Animal & Biodiversity	Cyber & Privacy Concerns	Economic & Consumer Rights	Environmental Issues	Health & Safety	Labor & Employment	Legal & Compliance	Political & Governance	Resource & Infrastructure	Social & Human Rights	Substance & Social Ethics	Technology & Surveillance	Weapons & Security
	Whole Sample	**0.082***	**0.022*****	0.009	0.008	-0.016	-0.001	**0.043*****	0.032	0.036	**0.038*****	**0.094****	0.176	0.032
		(0.048)	**(0.007)**	(0.022)	(0.008)	(0.013)	(0.017)	**(0.016)**	(0.041)	(0.036)	**(0.010)**	**(0.042)**	(0.366)	(0.031)
Political	Republican	-0.019	**0.032***	-0.005	-0.012	-0.022	-0.021	0.042	**0.424****	0.032	-0.028	0.145	0.008	0.108
		(0.091)	**(0.017)**	(0.041)	(0.010)	(0.017)	(0.043)	(0.040)	**(0.209)**	(0.050)	(0.051)	(0.089)	(0.628)	(0.091)
	Democratic	**0.249*****	**0.018****	0.020	**0.016***	-0.016	0.005	**0.049****	0.006	0.009	**0.048*****	0.053	0.417	0.013
		(0.077)	**(0.007)**	(0.028)	**(0.010)**	(0.020)	(0.019)	**(0.021)**	(0.056)	(0.056)	**(0.016)**	(0.049)	(0.446)	(0.032)
Flow Age	Old	0.083	**0.031****	0.053	-0.009	-0.018	0.014	0.030	0.076	-0.006	**0.050****	**0.115***	-0.018	-0.027
		(0.086)	**(0.012)**	(0.050)	(0.009)	(0.024)	(0.029)	(0.036)	(0.103)	(0.054)	**(0.024)**	**(0.059)**	(0.897)	(0.062)
	Young	0.064	**0.016***	-0.010	0.010	-0.016	-0.014	**0.042****	0.011	0.042	**0.030****	0.061	0.207	0.032
		(0.061)	(0.008)	(0.027)	(0.009)	(0.017)	(0.024)	**(0.019)**	(0.052)	(0.049)	**(0.012)**	(0.064)	(0.378)	(0.034)
TDF Age	Old	0.064	0.027	0.045	0.025	-0.015	0.007	0.024	-0.269	0.101	0.030	0.068	-0.802	-0.001
		(0.069)	(0.018)	(0.033)	(0.018)	(0.033)	(0.035)	(0.030)	(0.276)	(0.063)	(0.023)	(0.068)	(0.645)	(0.055)
	Young	**0.123***	**0.020*****	-0.006	0.004	-0.020	-0.010	**0.048****	0.039	0.018	**0.041*****	**0.112***	**0.640***	0.052
		(0.069)	**(0.008)**	(0.028)	(0.010)	(0.015)	(0.021)	**(0.020)**	(0.049)	(0.044)	**(0.012)**	**(0.058)**	**(0.388)**	(0.042)
Deferral	Poor	-0.171	0.045	-0.040	-0.085	-0.101	-0.014	**0.127***	0.169	0.162	-0.140	0.069	-0.516	**-0.404****
		(0.233)	(0.029)	(0.130)	(0.057)	(0.074)	(0.059)	**(0.077)**	(0.405)	(0.136)	(0.118)	(0.097)	(0.956)	**(0.197)**
	Rich	**0.118****	**0.023*****	-0.001	0.009	-0.019	-0.001	**0.033***	0.042	0.025	**0.033*****	**0.089***	0.331	0.017
		(0.058)	**(0.009)**	(0.020)	(0.008)	(0.014)	(0.019)	**(0.018)**	(0.054)	(0.042)	**(0.011)**	**(0.052)**	(0.412)	(0.031)
Account Balance	Poor	**0.266*****	0.014	0.024	0.037	0.001	-0.010	**0.073***	0.061	0.106	0.036	**0.189***	0.608	0.040
		(0.099)	(0.015)	(0.057)	(0.034)	(0.039)	(0.032)	**(0.038)**	(0.100)	(0.076)	(0.026)	**(0.097)**	(0.620)	(0.069)
	Rich	0.037	**0.028*****	0.003	0.001	-0.018	-0.002	**0.031***	0.010	0.010	**0.036*****	0.074	0.161	0.017
		(0.058)	**(0.009)**	(0.022)	(0.009)	(0.015)	(0.020)	**(0.018)**	(0.041)	(0.042)	**(0.011)**	(0.047)	(0.443)	(0.037)

Table 1.11: Impact of ESG scandals on ESG fund flows

The table presents the estimated coefficients for the fund-plan-year-level panel model $Flow\_Employee_{i,j,t} = \beta ESG\_Scandal_{i,t-1} + \mathbf{\Gamma} \boldsymbol{Controls_{i,j,t}} + \textbf{Fixed Effects}$. $ESG\_Scandal_{i,t-1}$ denotes the occurrence of ESG scandals in the first-two-digit zip code area of plan i during the preceding year $(t-1)$, and $Flow\_Employee_{i,j,t}$ represents the flow of fund j introduced by employees in plan i and year t. The variable $Controls_{i,j,t}$ includes the following covariates: log plan total assets ($Size$), log plan menu size ($MenuSize$), three-year return of fund j ($Ret_{t-1,t-3}$), and three-year monthly volatility ($Vol_{t-1,t-3}$). The regression incorporates year and plan indicators as control variables. The first two columns represent all fund-plan-year observations in the sample, distinguishing between ESG and non-ESG funds. The following columns focus on ESG funds. Subsamples are indicated in column headers. To account for potential within-plan correlation, standard errors are clustered at the plan level. Statistical significance levels are denoted by *, **, and ***, corresponding to 10%, 5%, and 1% significance levels, respectively.

	ESG Funds	Non-ESG Funds	Political		Flow Age		TDF Age		Deferral		Account Balance	
			Democratic	Republican	Old	Young	Old	Young	Poor	Rich	Poor	Rich
ESG_Scandal_t-1	**0.042****	**-0.018*****	-0.009	**0.050****	0.002	**0.036****	0.035	0.014	0.008	**0.032****	0.012	0.021
	(0.021)	**(0.005)**	(0.016)	**(0.020)**	(0.024)	**(0.016)**	(0.021)	(0.016)	(0.032)	**(0.016)**	(0.021)	(0.016)
PlanSize	0.054	0.487***	0.028	0.029	0.085	0.011	0.083*	-0.019	0.069	0.007	-0.072	0.051
	(0.065)	(0.020)	(0.057)	(0.041)	(0.053)	(0.044)	(0.044)	(0.056)	(0.081)	(0.053)	(0.062)	(0.044)
MenuSize	0.132***	0.141***	0.127***	0.116***	0.147***	0.109***	0.125***	0.145***	0.232***	0.129***	0.205***	0.128***
	(0.031)	(0.016)	(0.031)	(0.029)	(0.038)	(0.025)	(0.029)	(0.032)	(0.044)	(0.031)	(0.039)	(0.027)
Ret_t-1_t-3	0.080	0.217***	0.080	0.085**	0.083	0.087**	0.088**	0.083**	0.155**	0.065*	0.169***	0.053
	(0.059)	(0.013)	(0.052)	(0.038)	(0.052)	(0.040)	(0.044)	(0.042)	(0.071)	(0.039)	(0.065)	(0.037)
Vol_t-1-t-3	-0.051	-0.119***	-0.008	-0.040**	-0.067***	0.010	-0.053**	0.009	-0.020	-0.052***	0.005	-0.051***
	(0.031)	(0.003)	(0.025)	(0.019)	(0.023)	(0.021)	(0.022)	(0.020)	(0.039)	(0.020)	(0.035)	(0.019)
Observations	14,166	2,846,949	5,125	8,983	5,989	7,191	6,798	6,904	3,643	7,999	3,746	9,167
R-squared	0.248	0.062	0.252	0.270	0.315	0.278	0.299	0.268	0.295	0.266	0.291	0.275

Table 1.12: The priority role of political leanings in ESG fund flows

The table presents the estimated coefficients for the fund-plan-year-level panel model $Flow\_Employee_{i,j,t} = \beta ESG\_Scandal_{i,t-1} + \mathbf{\Gamma}\boldsymbol{Controls_{i,j,t}} +$ **Fixed Effects**. Only ESG funds are included in this table. Regression variables and model settings in the regressions are identical to Table 1.11. Each pair of coefficients and standard errors corresponds to a single regression conducted within specific subsamples, determined by a double-sorting mechanism based on presidential voting results and either wealth or age. The proxies for wealth and age follow the definitions outlined in Table 1.6. To address potential correlations, standard errors are clustered at the plan and year level. Statistical significance levels are denoted by *, **, and ***, corresponding to significance at the 10%, 5%, and 1% levels, respectively.

		Democratic		Republican	
		Coeff.	S.E.	Coeff.	S.E.
Flow Age	Old	0.034	(0.028)	0.031	(0.043)
	Young	-0.010	(0.024)	**0.072****	(0.032)
TDF Age	Old	0.042	(0.032)	0.044	(0.032)
	Young	-0.002	(0.021)	**0.029***	(0.014)
Deferral	Poor	-0.012	(0.032)	0.063	(0.057)
	Rich	0.003	(0.023)	**0.055****	(0.025)
Account Balance	Poor	-0.034	(0.032)	-0.010	(0.040)
	Rich	-0.011	(0.020)	**0.064*****	(0.025)

Table 1.13: ESG fund flow heterogeneity in responses to specific scandal issues

This table presents estimated coefficients for the fund-plan-year-level panel model $Flow\_Employee_{i,j,t} = \beta X_{i,t-1} + \mathbf{\Gamma}\boldsymbol{Controls_{i,j,t}} +$ **Fixed Effects**. The first three columns use $X_{i,t-1}$ to represent the number of social scandals ($Social_{i,t-1}$), environmental scandals ($Environ_{i,t-1}$), and governance scandals ($Gover_{i,t-1}$). In subsequent columns, $X_{i,t-1}$ refers to the occurrence of specific ESG scandals indicated in the column headings near plan i during the preceding year ($t-1$). The ESG scandals are categorized into 13 mutually exclusive categories from RepRisk's dataset. Each pair of coefficients and standard errors corresponds to a single regression conducted within specific subsamples indicated in the column headers. Other regression variables and model settings in the regressions are identical to Table 1.11. To address potential correlations, standard errors are clustered at the plan and year level. Statistical significance levels are denoted by *, **, and ***, indicating significance at the 10%, 5%, and 1% levels, respectively.

		$Social_{t-1}$	$Environ_{t-1}$	$Gover_{t-1}$	Animal & Biodiversity	Cyber & Privacy Concerns	Economic & Consumer Rights	Environmental Issues	Health & Safety	Labor & Employment	Legal & Compliance	Political & Governance	Resource & Infrastructure	Social & Human Rights	Substance & Social Ethics	Technology & Surveillance	Weapons & Security
Whole Sample		0.035**	0.069**	-0.008	0.003	0.000	0.004***	0.001**	0.001**	0.001	0.001	-0.003	0.002	0.001*	-0.001	-0.000	-0.002
		(0.014)	(0.028)	(0.011)	(0.003)	(0.000)	(0.001)	(0.000)	(0.001)	(0.001)	(0.001)	(0.003)	(0.002)	(0.001)	(0.002)	(0.016)	(0.002)
Political	Democratic	-0.000	0.050	-0.015	0.007	0.000	0.003*	0.001	0.002	-0.001	0.000	-0.002	0.002	-0.000	0.002	0.011	-0.004*
		(0.017)	(0.063)	(0.013)	(0.004)	(0.000)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)	(0.003)	(0.001)	(0.003)	(0.020)	(0.002)
	Republican	0.096***	0.066*	0.070	0.001	0.001	0.005**	0.001*	0.001	0.003*	0.001	-0.001	0.002	0.002*	-0.004	-0.017	0.000
		(0.029)	(0.034)	(0.046)	(0.004)	(0.001)	(0.002)	(0.000)	(0.001)	(0.002)	(0.002)	(0.009)	(0.003)	(0.001)	(0.002)	(0.029)	(0.005)
Flow Age	Old	0.020	0.069	-0.004	0.003	-0.001	0.003	0.001	0.001	0.003	0.000	-0.003	-0.002	-0.000	-0.003	-0.026	-0.003
		(0.031)	(0.051)	(0.023)	(0.005)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.008)	(0.004)	(0.001)	(0.005)	(0.031)	(0.004)
	Young	0.040**	0.064*	-0.003	0.005	0.001	0.003	0.001**	0.002**	-0.001	0.002	-0.002	0.004	0.001*	0.003	0.010	-0.002
		(0.017)	(0.037)	(0.016)	(0.004)	(0.001)	(0.002)	(0.000)	(0.001)	(0.001)	(0.001)	(0.004)	(0.003)	(0.001)	(0.002)	(0.022)	(0.003)
TDF Age	Old	0.060**	0.077	0.011	0.006	0.001*	0.002	0.001*	0.002	0.003*	0.000	-0.009	-0.002	0.001	-0.003	-0.003	-0.000
		(0.025)	(0.049)	(0.020)	(0.004)	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.006)	(0.003)	(0.001)	(0.002)	(0.027)	(0.003)
	Young	0.017	0.046	-0.018	0.003	0.000	0.004**	0.000	0.002	-0.000	0.001	-0.002	0.004	0.001*	0.002	0.011	-0.004
		(0.018)	(0.036)	(0.015)	(0.004)	(0.000)	(0.002)	(0.000)	(0.001)	(0.001)	(0.001)	(0.004)	(0.003)	(0.001)	(0.003)	(0.019)	(0.003)
Deferral	Poor	0.003	-0.049	-0.011	0.002	0.000	-0.001	-0.000	-0.003*	0.002	0.001	-0.015**	0.000	0.001	-0.004	0.034	0.000
		(0.032)	(0.069)	(0.064)	(0.007)	(0.001)	(0.004)	(0.001)	(0.002)	(0.002)	(0.002)	(0.007)	(0.004)	(0.001)	(0.004)	(0.039)	(0.005)
	Rich	0.048**	0.099***	-0.011	0.006*	0.001*	0.004***	0.001***	0.002**	0.000	0.002	0.002	0.003	0.002**	0.001	0.009	-0.002
		(0.019)	(0.037)	(0.017)	(0.003)	(0.000)	(0.002)	(0.000)	(0.001)	(0.001)	(0.001)	(0.005)	(0.003)	(0.001)	(0.003)	(0.020)	(0.002)
Account Balance	Poor	0.033	0.055	0.010	-0.001	-0.000	0.006**	0.001	0.001	0.000	0.002	0.002	0.007	0.002*	0.000	-0.012	0.001
		(0.028)	(0.045)	(0.023)	(0.006)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.009)	(0.005)	(0.001)	(0.004)	(0.035)	(0.004)
	Rich	0.034*	0.102***	-0.014	0.005*	0.001	0.004**	0.001***	0.002**	0.001	0.001	-0.002	0.001	0.001	-0.002	0.003	-0.002
		(0.018)	(0.039)	(0.015)	(0.003)	(0.000)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.004)	(0.002)	(0.001)	(0.002)	(0.019)	(0.002)

Table 1.14: Heckman sample selection in employee ESG fund flow

This table presents the results of a Heckman (1979) self-selection model examining the impact of ESG scandals on ESG fund flows. The first-stage regression includes $ESG\_Scandal_{i,t-1}$ as the independent variable along with control variables: log plan size ($Size$), log menu size ($MenuSize$), an indicator for employee auto-enrollment ($AutoEnroll$), an indicator for target-date funds (TDFs) as the default investment option ($DefaultInv$), an indicator for the mutual fund company serving as the plan's main service provider ($DuralRole$), plan indicators, and year indicators. In the second-stage regression, the same control variables as in Table 1.11 are used. The lower panel displays the first-stage regression coefficients, D, of $ESG\_Scandal_{i,t-1}$, while the upper panel reports the direct impact of $ESG\_Scandal_{i,t-1}$ on fund flows and the estimated coefficients of Heckman (1979)'s λ. Each column represents a single regression. The subsamples are indicated in the column headers. To address sampling issues, standard errors are bootstrapped. Statistical significance levels are denoted by *, **, and ***, indicating significance at the 10%, 5%, and 1% levels, respectively.

		Political		Flow Age		TDF Age		Deferral		Account Balance	
	Whole Sample	Democratic	Republican	Old	Young	Old	Young	Poor	Rich	Poor	Rich
$ESG\_Scandal_{t-1}$	0.017	0.004	0.031	0.035	0.007	0.048	-0.007	0.046	0.001	-0.010	0.024
	(0.020)	(0.027)	(0.031)	(0.035)	(0.022)	(0.032)	(0.024)	(0.045)	(0.025)	(0.037)	(0.025)
λ	**-0.287*** **	**-0.379*** **	**-0.260*** **	**-0.508*** **	**-0.107** **	**-0.394*** **	**-0.198*** **	**-0.420** **	**-0.253*** **	-0.212	**-0.286*** **
	(0.057)	**(0.122)**	**(0.064)**	**(0.135)**	**(0.045)**	**(0.119)**	**(0.053)**	**(0.174)**	**(0.084)**	(0.150)	**(0.080)**
D	**0.043*** **	**0.044** **	0.015	**0.042** **	**0.037** **	**0.032* **	**0.045** **	-0.010	**0.076*** **	**0.056** **	**0.052*** **
	(0.014)	**(0.019)**	(0.021)	**(0.021)**	**(0.019)**	**(0.018)**	**(0.019)**	(0.028)	**(0.018)**	**(0.027)**	**(0.017)**
Observations	108,331	33,877	74,454	48,921	59,410	57,832	50,494	35,548	61,161	28,950	74,264

Table 1.15: Heterogeneous investor preferences among specific ESG scores in fund additions

The table presents the estimated coefficients for the plan-year-level panel model $Y_{i,t} = \boldsymbol{\beta X_{i,t-1}} + \boldsymbol{\Gamma Controls_{i,t}} + \textbf{Fixed Effects}$, where the dependent variables $Y_{i,t}$ are the specific environment, social, and governance scores of ESG funds added to plan i in year t. The scores are set to zero for all plans without ESG fund additions. The specific scores are demeaned within each Morningstar category in the last three columns. $\boldsymbol{X_{i,t}}$ contains three independent variables, including social, environmental and governance scandals in $t-1$ near plan i. Each pair of coefficient and standard error represents a single regression. Each panel indicates the subsample determination at the panel heads. All other variable settings are identical to Table 1.3. All regressions include year and plan fixed effects as well. Standard errors are clustered at the plan and year level to account for potential correlations. Statistical significance levels are denoted by *, **, and ***, corresponding to the 10%, 5%, and 1% significance levels, respectively.

	E-Score	S-Score	G-Score		E-Score	S-Score	G-Score
		Whole Sample					
Social	**0.193*** **	**0.191*** **	0.054				
	(0.047)	**(0.050)**	(0.046)				
Environmental	**0.120***	0.018	0.070				
	(0.068)	(0.104)	(0.117)				
Governance	**-0.063** **	-0.018	0.012				
	(0.026)	(0.022)	(0.032)				
		Republican				Democratic	
Social	0.151	**0.196*** **	0.035	Social	**0.150*** **	**0.145** **	0.060
	(0.086)	**(0.054)**	(0.065)		**(0.043)**	**(0.044)**	(0.060)
Environmental	-0.534	0.063	0.140	Environmental	**0.292** **	-0.152	-0.105
	(0.329)	(0.120)	(0.150)		**(0.108)**	(0.173)	(0.293)
Governance	-0.119	-0.060	-0.118	Governance	-0.050**	0.006	0.023
	(0.092)	(0.045)	(0.100)		(0.018)	(0.022)	(0.041)
		Young				Old	
Social	**0.220*** **	**0.240*** **	0.090	Social	0.098	**0.131** **	0.017
	(0.061)	**(0.047)**	(0.055)		(0.073)	**(0.051)**	(0.049)
Environmental	**0.106** **	0.093	0.261*	Environmental	0.031	-0.141	-0.181
	(0.057)	(0.222)	(0.118)		(0.111)	(0.154)	(0.196)
Governance	**-0.080***	-0.042	0.012	Governance	-0.028	0.031	0.024
	(0.035)	(0.031)	(0.027)		(0.027)	(0.034)	(0.050)
		Poor				Rich	
Social	0.129	**0.218** **	0.051	Social	**0.198*** **	**0.171** **	0.058
	(0.107)	**(0.105)**	(0.090)		**(0.045)**	**(0.055)**	(0.041)
Environmental	0.452	0.484	0.304	Environmental	-0.108	-0.158	0.013
	(0.460)	(0.291)	(0.176)		(0.164)	(0.133)	(0.136)
Governance	-0.049	-0.062	0.034	Governance	**-0.080*** **	-0.022	0.010
	(0.088)	(0.062)	(0.051)		**(0.015)**	(0.023)	(0.035)

Table 1.16: Heterogeneous investor preferences among specific ESG scores in fund flows

This table presents estimated coefficients $\boldsymbol{\beta}$s for the fund-plan-year-level panel model:

$$
\begin{aligned}
Flow\_Employee_{i,j,t} &= \beta_1 E\_Score_{j,t} \times ESG\_Scandal_{i,t-1} + \beta_2 S\_Score_{j,t} \times ESG\_Scandal_{i,t-1} \\
&\quad + \beta_3 G\_Score_{j,t} \times ESG\_Scandal_{i,t-1} + \beta_0 ESG\_Scandal'_{i,t-1} \\
&\quad + \boldsymbol{\Gamma\, Controls_{i,j,t}} + \textbf{Fixed Effects}
\end{aligned}
$$

where $E\_Score_{j,t}$, $S\_Score_{j,t}$, and $G\_Score_{j,t}$ are the specific environment, social, and governance scores of ESG funds added to plan i in year t. The estimated coefficients β_i where $i \in (1, 2, 3)$ are shown in the table. To mitigate potential multicollinearity, we use the logarithm of $ESG\_Scandal_{i,t-1}$ denoted as $ESG\_Scandal'_{i,t-1}$. All the other variable settings are identical to Table 1.11. Each column represents a single regression, and the subsample determinations are shown in the column headers. To address potential correlations, standard errors are clustered at the plan and year level. Statistical significance levels are denoted by *, **, and ***, indicating significance at the 10%, 5%, and 1% levels, respectively.

	Whole Sample	Political		TDF Age		Flow Age		Deferral		Account Balance	
		Democratic	Republican	Young	Old	Young	Old	Poor	Rich	Poor	Rich
β_1	0.003	0.006	**0.011***	0.007	-0.004	0.002	-0.001	-0.000	0.006	-0.005	0.005
	(0.003)	(0.004)	**(0.006)**	(0.004)	(0.005)	(0.004)	(0.006)	(0.008)	(0.004)	(0.005)	(0.004)
β_2	**0.018*****	0.011	**0.029*****	**0.025*****	0.006	**0.024****	0.005	0.008	**0.019*****	0.016	**0.012***
	(0.006)	(0.008)	**(0.010)**	**(0.009)**	(0.009)	**(0.009)**	(0.010)	(0.014)	**(0.007)**	(0.012)	**(0.007)**
β_3	**-0.012*****	**-0.018*****	**-0.016***	**-0.017*****	0.002	**-0.018*****	-0.005	-0.009	**-0.019*****	-0.006	**-0.015*****
	(0.005)	**(0.006)**	**(0.009)**	**(0.006)**	(0.007)	**(0.006)**	(0.010)	(0.012)	**(0.006)**	(0.010)	**(0.005)**

1.A Appendix

1.A.1 ESG Indices Dispersion

Figure 1.A.1 illustrates a persistent absence of substantial correlations among different ESG rating providers over the last decade, with Asset4 and MSCI ratings notably displaying a sub-0.3 correlation. These obstructions in finding a proper ESG portfolio and fulfilling the non-pecuniary demand could lead to sub-optimal equilibrium (Hartzmark and Sussman, 2019).

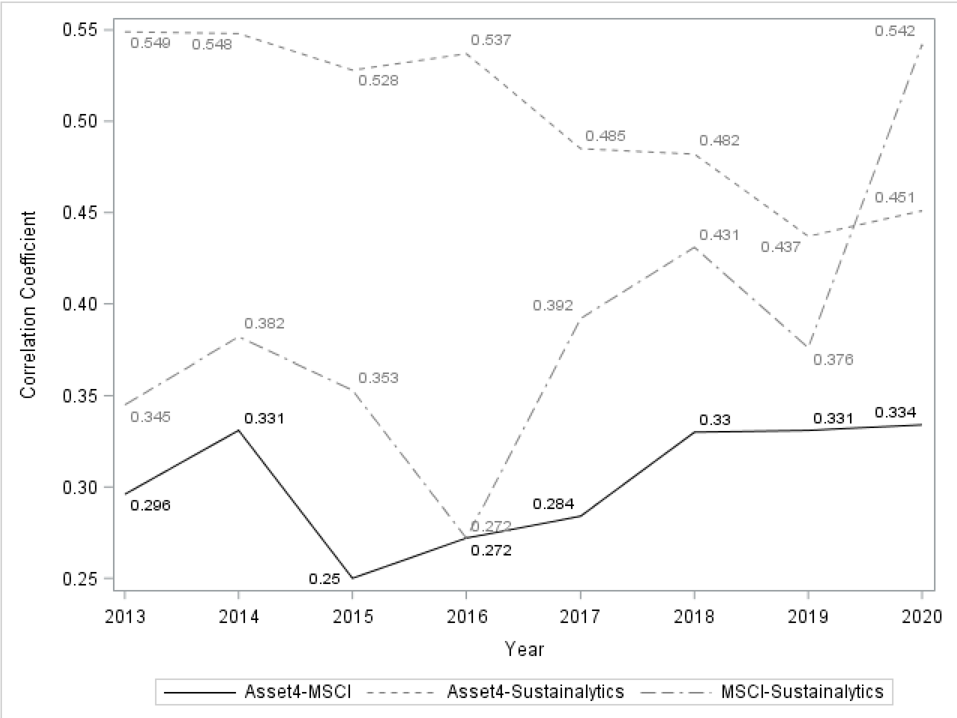


Figure 1.A.1: Pairwise correlations of ESG scores

1.A.2 Categories of ESG Scandals

In Table 1.A.1, all 73 specific tags defined by RepRisk are categorized into 13 issues on our own.

Table 1.A.1: Category of scandal issues

Issue Name	Category
Animal transportation	Animal & Biodiversity
Endangered species	Animal & Biodiversity
Fur and exotic animal skins	Animal & Biodiversity
Cyberattack	Cyber & Privacy Concerns
Privacy violations	Cyber & Privacy Concerns
Access to products and service	Economic & Consumer Rights
Agricultural commodity speculation	Economic & Consumer Rights
Conflict minerals	Economic & Consumer Rights
Diamonds	Economic & Consumer Rights
Economic impact	Economic & Consumer Rights
Land grabbing	Economic & Consumer Rights
Predatory lending	Economic & Consumer Rights
Rare earths	Economic & Consumer Rights
Tax havens	Economic & Consumer Rights
Abusive/Illegal fishing	Environmental Issues
Airborne pollutants	Environmental Issues
Arctic drilling	Environmental Issues
Asbestos	Environmental Issues
Coal-fired power plants	Environmental Issues
Coral reefs	Environmental Issues
Deep sea drilling	Environmental Issues
Forest burning	Environmental Issues
Fracking	Environmental Issues
Greenhouse gas (GHG) emissions	Environmental Issues
High conservation value forests	Environmental Issues
Illegal logging	Environmental Issues
Land ecosystems	Environmental Issues
Marine/Coastal ecosystems	Environmental Issues
Monocultures	Environmental Issues
Mountaintop removal mining	Environmental Issues
Offshore drilling	Environmental Issues
Oil sands	Environmental Issues
Palm oil	Environmental Issues
Plastics	Environmental Issues
Protected areas	Environmental Issues
Sand mining and dredging	Environmental Issues
Seabed mining	Environmental Issues

Table 1.A.1: Category of scandal issues

Issue Name	Category
Ship breaking and scrapping	Environmental Issues
Soy	Environmental Issues
Epidemics/Pandemics	Health & Safety
Genetically modified organisms (GMOs)	Health & Safety
Health impact	Health & Safety
Migrant labor	Labor & Employment
Salaries and benefits	Labor & Employment
Negligence	Legal & Compliance
Security services	Legal & Compliance
Lobbying	Political & Governance
Energy management	Resource & Infrastructure
Hydropower (dams)	Resource & Infrastructure
Nuclear power	Resource & Infrastructure
Wastewater management	Resource & Infrastructure
Water management	Resource & Infrastructure
Water scarcity	Resource & Infrastructure
Gender inequality	Social & Human Rights
Genocide/Ethnic cleansing	Social & Human Rights
Human trafficking	Social & Human Rights
Indigenous people	Social & Human Rights
Involuntary resettlement	Social & Human Rights
Racism/Racial inequality	Social & Human Rights
Alcohol	Substance & Social Ethics
Gambling	Substance & Social Ethics
Marijuana/Cannabis	Substance & Social Ethics
Opioids	Substance & Social Ethics
Pornography	Substance & Social Ethics
Tobacco	Substance & Social Ethics
Drones	Technology & Surveillance
Automatic and semi-automatic weapons	Weapons & Security
Biological weapons	Weapons & Security
Chemical weapons	Weapons & Security
Cluster munitions	Weapons & Security
Depleted uranium munitions	Weapons & Security
Land mines	Weapons & Security
Nuclear weapons	Weapons & Security

1.A.3 Mechanism of Scandal Impact in Portfolio Choice

There is much anecdotal evidence for the direct impact of ESG scandals on 401(k) investment decisions of nearby plans. Figure 1.A.2 shows two scandals of two firms headquartered in Philadelphia in 2012. This year, the number of scandals that happened in Philadelphia increased from 12 in 2011 to 26. In response, more than ten plans in this area added ESG funds to their menus in 2013. Figure 1.A.3 shows a company's 401(k) investment menu in 2013. There are three ESG funds in the menu, which are all newly added.

The challenges to the scandal impact setting on ESG investment come from two parts: the concern of multi-branches firms and the intrinsic difference of geographic locations. Due to the data limitation, we assume the scandals have the most significant effect on investment around the scandal firm's headquarters location. The scandal could happen around any branch and even outside any of the branches. For example, an oil spill happens in the ocean. This is only from the real-impact part. The scandals have two parts in our analysis: real impact and psychological impact. The headquarters is assumed to have the strongest impact of adding these two parts together. The geographic locations are, therefore, another source of challenges. The variation of the number of scandals is intrinsically higher for a less populated area, for example, Nevada, given one scandal happens. However, it is quite low for the high population density areas, for example, California. The intrinsic differences between these two areas, including political leanings, culture and education, could all lead to outlier results of our analysis.

To relieve the first challenge, we change the independent variable to $Scandal\_Single_{i,t-1}$ as the occurrence of scandals from firms with only one branch. Although the scandals could still happen outside the firm's location, the impact of the scandals is more direct to the local residents. In the first three columns of Table 1.A.2, we run regressions following the same setting in the menu-change case, but with $Scandal\_Single_{t-1}$ substituting $ESG\_Scandal_{t-1}$. The coefficients are still significant for the conditional on menu-change case, with even a higher magnitude of point estimation compared with main results. However, in the unconditional case, the coefficient is insignificant. This is because by using the single branch firm, we actually abandon large companies, and these companies usually have profound impacts on the local residents.

We use the unexpected scandals in column 4 to 6 of Table 1.A.2 to partially cancel out the intrinsic time-series differences of geographic location. The variable $Scandal\_Detrend_{i,t-1}$ is the **positive** part of

$$ESG\_Scandal_{i,t-1} - \frac{1}{5} \sum_{t-6 \leq j \leq t-2} ESG\_Scandal_{i,j},$$

and zero for the negative value of the above expression. This variable captures the unexpected happening of the scandals in a given area. All the coefficients show high significance here. Although this could not cancel all time-series differences in geographic location, by combining the plan and location fixed effect, we have controlled most of it. The magnitude of coefficients is lower than the original case, which could be attributed to the intrinsic differences among areas. This is a support for the challenge, but results still hold.

The psychological reason for investing in ESG is probably not return chasing or expectation changes. In Table 1.A.3, we use the future alphas of added ESG funds as the dependent variable. All insignificant coefficients show that the variation in ESG scandals nearby does not link to a higher return-chasing propensity. Further, in Table 1.A.4, it is also insignificant that such ESG scandals could lead to investing in more high-return products. If we treat realized return as expected return, investors show no evidence in return-chasing concerns when responding to ESG scandals.

1.A.4 Static Differences in ESG Fund Preference

To account for additional plan-level factors, we perform the regression analysis presented in Table 1.A.5. The results demonstrate that political leaning and age exhibit the strongest and most significant correlations with ESG investment in 401(k) plans. We define the dependent variable as the ratio of ESG funds to the total number of investment options within each plan and follow the settings in Table 1.3.

(a) Bribery scandal of FMC Corp.

HOME » COMCAST/XFINITY »CONSUMER NEWS »VIDEO » CURRENTLY READING:

Comcast Mistakenly Switches Good Morning America With Hardcore Porn in Colorado Springs

Phillip Dampier | April 18, 2012 | Comcast/Xfinity, Consumer News, Video | No Comments

Comcast Cable subscribers in Colorado Springs got more than a bowl of Froot Loops Thursday morning when local ABC affiliate KRDO was suddenly replaced with hardcore adult pornography during an airing of *Good Morning America*.

(b) Pornography scandal of Comcast

Figure 1.A.2: Scandals happened around Philadelphia

Franklin-IncomeFund	Mutual fund	**	101,139
American Century Disciplined Growth Fund	Mutual fund	**	73,074
Oppenheimer-International Growth Fund	Mutual fund	**	50,189
Vanguard Total Stock Market Index Fund	Mutual fund	**	45,056
Alliance-Small Cap Growth Fund	Mutual fund	**	41,613
Neuberger-Socially Responsive Fund	Mutual fund	**	34,339
Loomis-Bond Fund	Mutual fund	**	33,730
Oppenheimer-Developing Markets Fund	Mutual fund	**	31,504
Loomis-Investment Grade Fund	Mutual fund	**	31,412
Transamerica-Ivy Science Fund	Mutual fund	**	31,247
Alger-Green Fund	Mutual fund	**	30,829
Prudential-Mid Cap growth Fund	Mutual fund	**	29,750
Vanguard REIT Index Fund	Mutual fund	**	24,270
Black Rock Health Science Fund	Mutual fund	**	24,117
Wells Fargo-Govt. Securities Fund	Mutual fund	**	22,033
American Fund-New Perspective Fund	Mutual fund	**	18,661
Vanguard Small Cap Index Fund	Mutual fund	**	18,154
Vanguard Target Retirement 2050 Fund	Mutual fund	**	13,543
American Fund-Small Cap World Fund	Mutual fund	**	10,949
Templeton-Global Bond Fund	Mutual fund	**	9,932
Vanguard Total International Stock Index Fund	Mutual fund	**	9,320

Figure 1.A.3: Greater Philadelphia Action responses to the scandals

In the first column of Table 1.A.5, ESG scores of nearby public firms (Geo_ESG) do not yield significant explanatory power for the variation in ESG investment. Conversely, the second column reveals that a 1% increase in Democratic Party votes results in a 6-percentage-point higher proportion of ESG funds within the 401(k) menu. Although the proxy for participants' age (flow) is statistically insignificant, the TDF age proxy yields significance in the third column. Specifically, a 1% decrease in the average age of participants corresponds to a 13-percentage-point higher allocation of ESG funds in the menu. The analysis does not uncover a significant relationship between wealth and ESG investment, as evidenced by the non-significant coefficients in columns 5 and 6.

Table 1.A.2: Robustness check: scandal settings

The following table presents the estimated coefficients for the plan-year level panel model similar to Table 1.3. In the first panel, the variable $ESG\_Scandals$ is substituted by $Scandal\_Single_{i,t-1}$ or $Scandal\_Detrend_{i,t-1}$. $Scandal\_Single_{i,t-1}$ is the occurrence of scandals from firms with only one branch in year t near plan i. $Scandal\_Detrend_{i,t-1}$ is the positive part of the difference between $ESG\_Scandals_{i,t-1}$ and past five-year average of $ESG\_Scandals_{i,j}$ with $j \in (t-6, t-2)$. The regression models are indicated at the column heads and the conditional cases are indicated at the bottom of each column. All regressions include plan and year fixed effects. All standard errors are clustered at the plan and year level to account for potential correlations. Levels of statistical significance are denoted by *, **, and ***, corresponding to the 10%, 5%, and 1% significance levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
Reg Method	OLS	Logit	Logit	OLS	Logit	Logit
$Scandal\_Single_{t-1}$	0.026**	0.345	1.038*			
	(0.012)	(0.265)	(0.602)			
$Scandal\_Detrend_{t-1}$				0.009*	0.165***	0.283*
				(0.005)	(0.056)	(0.159)
Conditional	Change	Uncon	Change	Change	Uncon	Change

Table 1.A.3: Return concern on ESG fund addition

The table presents the estimated coefficients (βs) for the plan-year-level panel model: $y_{i,t} = \beta ESG\_Scandal_{i,t-1} + \mathbf{\Gamma} \boldsymbol{Controls_{i,t}} + \textbf{Fixed Effects}$, where $y_{i,t}$ includes the expected αs corresponding to additional funds in the following year $t+1$. In the left panel, as shown in the top bar, the $y_{i,t}$s represent the added ESG funds' αs, while the right panel includes all added funds' αs. All other settings, including variable definitions and table structure, are identical to Table 1.7. To account for potential correlations, standard errors are clustered at the plan and year level. Statistical significance levels are denoted by *, **, and ***, corresponding to the 10%, 5%, and 1% significance levels, respectively.

		ESG Fund Addition				All Fund Addition			
		Equal Weighted		Value Weighted		Equal Weighted		Value Weighted	
Whole Sample		-0.002	(0.004)	-0.002	(0.005)	-0.007	(0.008)	-0.009	(0.009)
Political Leaning	Democratic	-0.004	(0.006)	-0.003	(0.006)	-0.013	(0.012)	-0.015	(0.012)
	Republican	0.007	(0.007)	0.006	(0.007)	-0.000	(0.013)	-0.002	(0.014)
TDF Age	Old	-0.003	(0.009)	-0.003	(0.009)	0.005	(0.013)	0.005	(0.013)
	Young	-0.001	(0.005)	-0.001	(0.004)	-0.016	(0.013)	-0.020	(0.013)
Flow Age	Old	0.007	(0.009)	0.008	(0.009)	0.005	(0.013)	0.003	(0.014)
	Young	-0.011	(0.015)	-0.012	(0.015)	-0.014	(0.013)	-0.016	(0.014)
Deferral	Poor	-0.006	(0.006)	-0.007	(0.006)	0.011	(0.015)	0.011	(0.016)
	Rich	0.002	(0.007)	0.002	(0.007)	-0.013	(0.011)	-0.017	(0.012)
Account Balance	Poor	-0.006	(0.008)	-0.005	(0.009)	-0.015	(0.017)	-0.017	(0.017)
	Rich	-0.001	(0.006)	-0.002	(0.006)	-0.003	(0.010)	-0.004	(0.010)

Table 1.A.4: Preference in ESG Scores in ESG Funds

This table presents estimated coefficients $\boldsymbol{\beta}$s for the fund-plan-year level panel model:

$$\begin{aligned} Flow\_Employee_{i,j,t} &= \beta_1 E[\alpha_{j,t+1}] \times ESG\_Scandal_{i,t-1} + \beta_2 ESG\_Scandal_{i,t-1} + \beta_3 E[\alpha]_{j,t+1} \\ &\quad + \boldsymbol{\Gamma Controls_{i,j,t}} + \textbf{Fixed Effects} \end{aligned}$$

where $E[\alpha]_{j,t+1}]$ is the realized return of ESG fund j in plan i in year $t+1$. The estimated coefficients β_i where $i \in (1,2,3)$ are shown in the table. All the other variable settings are identical to Table 1.11. Each column represents a single regression, and the subsample determinations are shown at the top of each column. To address potential correlations, standard errors are clustered at the plan and year level. Statistical significance levels are denoted by *, **, and ***, indicating significance at the 10%, 5%, and 1% levels, respectively.

VARIABLES	Whole Sample	Political		TDF Age		Flow Age		Deferral		Account Balance	
		Democratic	Republican	Young	Old	Young	Old	Poor	Rich	Poor	Rich
β_1	0.000	0.000	0.000	-0.000	0.000	-0.000	0.000	0.000	0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)	(0.000)
β_2	0.045**	0.020	0.069**	0.063*	0.029	0.062**	0.011	0.013	0.061**	0.021	0.049*
	(0.019)	(0.028)	(0.032)	(0.034)	(0.027)	(0.026)	(0.038)	(0.050)	(0.026)	(0.031)	(0.025)
β_3	0.001	0.001	0.000	0.002	0.000	0.001	-0.000	-0.000	0.001	-0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)

1.A.5 A Model: ESG Benchmark and ESG Premium with Heterogeneous ESG Preferences

We discovered that investors' reactions to ESG (Environmental, Social, and Governance) incidents and externalities tend to prioritize social and environmental factors. Additionally, their investments allocate 50% more towards the social component and 30% towards the environmental aspect of ESG funds. However, it's noteworthy that ESG funds predominantly adhere to a 1:1:1 assessment system, despite these observed discrepancies in investor preferences. Pavlova and Sikorskaya (2023) and Kashyap et al. (2023) provide the ESG fund manager's investing choice with the concern of benchmark. Pástor, Stambaugh and Taylor (2021) models the heterogeneous preference and its impact on ESG premium.

Two Types of Investors and One Mutual Fund

There are three types of participants: direct investors (D), fund managers (M) and fund investors (F), with the initial wealth ratio of λ_D, 0 and λ_F to the total wealth of the world W_0. All the investors are mean-variance investors and continuum with an initial endowment of 1. There are two periods $0, 1$, and N risky stocks with the stochastic return of $r = \mu + \varepsilon$ and $\varepsilon \sim N(0, \Sigma)$. The risk stocks have a loading on ESG-relevant elements, L, a $N \times S$ matrix which we have S elements on ESG assessments. The total supply of all stocks is all one.

Direct investors choose the portfolio weights X_D to maximize the expected exponential utility of

$$\mathbf{E}\left[-\exp(-\gamma_D(1 + X_D'r) - X_D'LG_D)\right] \tag{1.1}$$

where G_D is the ESG evaluation on the utility of investor D. The F.O.C gives that

$$X_D = \frac{1}{\gamma_D}\Sigma^{-1}(\mu + \frac{1}{\gamma_D}LG_D). \tag{1.2}$$

The compensation to the fund manager is given by

$$w_M = \hat{a}X_M'r + b(X_M'r - \bar{X}r) + c(X_M'LG_F^* - \bar{X}LG_F^*) + \phi \tag{1.3}$$

Table 1.A.5: Plan characteristics and ESG fund coverage

The table presents the estimated coefficients for the plan-year level panel model $ESG\_Fund_{i,t} = \beta X_{i,t} + \mathbf{\Gamma} \boldsymbol{Controls_{i,t}} + \textbf{Fixed Effects}$, where $ESG\_Fund_{i,t}$ represents the percentage of ESG funds in plan i at year t, and $X_{i,t-1}$ denotes the plan characteristics in the preceding year ($t-1$). The variable $Controls_{i,t}$ includes the following covariates: log plan size ($Size$), log menu size ($MenuSize$), an indicator for employee auto-enrollment ($AutoEnroll$), an indicator for TDF as the default investment option ($DefaultInv$), and an indicator for the mutual fund company serving as the plan's main service provider ($DuralRole$). Moreover, the independent variables in the model comprise the following: R_Vote, which represents the presidential voting share for the Republican Party in the first-two-digit area; D_Vote, which denotes the presidential voting share for the Democratic Party; Geo_ESG, which represents the value-weighted ESG scores of all public companies in the first-two-digit area; $Flow_{TDF \leq 2040}/Flow_{TDF>2040}$ the ratio between the total flow to target-date funds (TDFs) no later than 2040 and the total flow to TDFs later than 2040; Log_Age, the logarithm of value-weighted TDF-based age; $Deferral$, the logarithm of the average employee contribution; and $Account_Balance$, the logarithm of the average account balance. Additionally, the table includes year and plan dummies as control variables. The standard errors are clustered at the plan and year level to account for potential correlations. Statistical significance levels are denoted by *, **, and ***, corresponding to the 10%, 5%, and 1% significance levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)
R_Vote		0.015				
		(0.024)				
D_Vote		**0.059*****				
		(0.023)				
Geo_ESG	-0.002					
	(0.002)					
$Flow_{TDF \leq 2040}/Flow_{TDF>2040}$			0.000			
			(0.000)			
Log_Age				**-0.132*****		
				(0.040)		
Deferral					0.084	
					(0.081)	
Account_Balance						0.008
						(0.007)
Observations	116,662	118,905	85,165	93,397	113,239	105,488
R-squared	0.757	0.756	0.810	0.800	0.776	0.774

Table 1.A.6: Unconditional leaning to E/S/G components

The table presents the unconditional leaning to E, S and G components when investors make investment decisions in 401(k). The learnings in the first two columns are the weighted average sub-scores in E, S and G discussed in Table 1.16 in every subsample portfolio. The political leaning, age and wealth sub-sample divisions follow Table 1.15. The differences between the first and the second column are reported in the third column and the t-stats are reported in the last column.

	Political Leaning			
	Republican	Democratic	Diff	t-stat
E-Score	-0.107	0.233	**-0.340**	-15.309
S-Score	-0.193	-0.173	**-0.020**	-2.091
G-Score	0.020	0.069	**-0.048**	-4.572
	Age			
	Young	Old	Diff	t-stat
E-Score	0.088	-0.063	**0.151**	7.044
S-Score	-0.223	-0.146	**-0.077**	-8.605
G-Score	0.001	0.077	**-0.076**	-7.446
	Wealth			
	Poor	Rich	Diff	t-stat
E-Score	-0.065	-0.015	**-0.050**	-2.040
S-Score	-0.199	-0.203	0.005	0.441
G-Score	0.027	0.017	0.010	0.857

Table 1.A.7: Heterogeneous investor preferences among specific ESG scores in fund additions

The table presents the estimated coefficients for the plan-year-level panel model $Y_{i,t} = \beta ESG\_Scandal_{i,t-1} + \mathbf{\Gamma}\boldsymbol{Controls_{i,t}} + \textbf{Fixed Effects}$, where the dependent variables $Y_{i,t}$ are the specific environment, social, and governance scores of ESG funds added to plan i in year t. The scores are set to zero for all plans without ESG fund additions. The specific scores are presented as raw numbers in the first three columns and demeaned within each Morningstar category in the last three columns. Each pair of coefficient and standard error represents a single regression. The left bar in each row indicates the subsample determination. All other variable settings are identical to Table 1.3. All regressions include year and plan fixed effects as well. Standard errors are clustered at the plan and year level to account for potential correlations. Statistical significance levels are denoted by *, **, and ***, corresponding to the 10%, 5%, and 1% significance levels, respectively.

		Raw			Demean		
		E-Score	S-Score	G-Score	E-Score	S-Score	G-Score
Whole Sample		**0.009****	**0.007****	**0.007***	**0.046***	**0.044****	**0.036****
		(0.004)	**(0.003)**	**(0.004)**	**(0.024)**	**(0.022)**	**(0.018)**
Political	Democratic	**0.023*****	**0.020*****	**0.020*****	**0.046**	**0.103****	**0.076****
		(0.008)	**(0.006)**	**(0.007)**	**(0.042)**	**(0.044)**	**(0.033)**
	Republican	0.009	0.006	0.005	**0.086****	0.040	0.005
		(0.006)	(0.005)	(0.006)	**(0.043)**	(0.034)	(0.031)
TDF Age	Young	**0.022*****	**0.017*****	**0.017****	0.071	**0.085****	0.049
		(0.007)	**(0.006)**	**(0.007)**	(0.047)	**(0.043)**	(0.034)
	Old	**0.014*****	**0.013****	**0.014****	**0.062****	**0.083****	0.033
		(0.006)	**(0.005)**	**(0.006)**	**(0.031)**	**(0.035)**	(0.027)
Flow Age	Young	**0.021*****	**0.016****	**0.017****	**0.090***	**0.100****	0.057
		(0.008)	**(0.006)**	**(0.007)**	**(0.046)**	**(0.043)**	(0.035)
	Old	**0.015*****	**0.014*****	**0.014****	0.042	**0.062****	0.029
		(0.006)	**(0.005)**	**(0.006)**	(0.033)	**(0.030)**	(0.026)
Deferral	Poor	**0.018****	**0.014****	**0.015***	0.035	0.030	0.057
		(0.008)	**(0.007)**	**(0.008)**	(0.043)	(0.041)	(0.037)
	Rich	**0.017*****	**0.013****	**0.013****	**0.047***	**0.081****	0.047
		(0.006)	**(0.005)**	**(0.006)**	**(0.025)**	**(0.036)**	(0.030)
Account Balance	Poor	0.013	0.010	0.009	0.085	0.057	0.050
		(0.009)	(0.008)	(0.009)	(0.065)	(0.060)	(0.041)
	Rich	**0.018*****	**0.014*****	**0.015*****	0.047	**0.077****	0.039
		(0.006)	**(0.005)**	**(0.005)**	(0.032)	**(0.031)**	(0.027)

where $\bar{X}$ is the market weight of each stock and G_F^* is the ESG evaluation on the social norm. Let's first assume the contract $(\hat{a}, b, c, \phi)$ are a given one to all managers. A fund manager's utility is just the mean-variance and depends only on the expected compensation. So, the optimization problem faced by fund managers is

$$\max_{X_M} \mathbf{E}\left[-\exp(-\gamma_M w_M)\right] = \max_{X_M} \mathbf{E}\left[-\exp(-\gamma_M(aX_M^{'}r - b\bar{X}r + cX_M^{'}LG_F^*))\right] \tag{1.4}$$

where $a = \hat{a} + b$. The F.O.C gives that

$$X_m = \frac{1}{a^2\gamma_M}\Sigma^{-1}(a\mu + cLG_F^*) + \frac{b}{a}\bar{X}. \tag{1.5}$$

Assume fund investors only invest to fund managers with all the wealth $\lambda_F W_0$, so the market weight (derived from market clearing) is given by

$$\begin{aligned}
\bar{X} &= \int_i \omega_i X_i di \\
&= \lambda_D X_D + \lambda_F X_M \\
&= \frac{\lambda_D}{\gamma_D}\Sigma^{-1}(\mu + \frac{1}{\gamma_D}LG_D) + \frac{\lambda_F}{a^2\gamma_M}\Sigma^{-1}(a\mu + cLG_F^*) + \frac{\lambda_F b}{a}\bar{X} \\
&\Longleftrightarrow \\
\bar{X} &= \frac{a}{a-\lambda_F}\left[\underbrace{\left(\frac{\lambda_D}{\gamma_D} + \frac{\lambda_F}{a\gamma_M}\right)\Sigma^{-1}\mu}_{\substack{\text{Average mean-variance} \\ \text{trade-off}}} + \underbrace{\frac{\lambda_D}{\gamma_D^2}\Sigma^{-1}LG_D}_{\substack{\text{ESG premium} \\ \text{from direct investor}}} + \underbrace{\frac{c\lambda_F}{a^2\gamma_M}\Sigma^{-1}LG_F^*}_{\substack{\text{ESG Premium} \\ \text{from fund manager/investor}}}\right].
\end{aligned}$$

The key thing is fund investors' G_F is different from the fund managers' G_F^*. G_F^* is from a social norm that for example if $S = 3$ for Environmental, Social, and Governance respectively, $G_{i,F}^* = (1, 1, 1)$ for the every $i = 0, 1, \ldots, N$. However, we have found that G_F would overweight environment and social part. If the fund investors could invest directly and have the perfect information on L, then their optimal investment should be

$$X_F^* = \frac{1}{\gamma_F}\Sigma^{-1}(\mu + \frac{1}{\gamma_F}LG_F). \tag{1.6}$$

but they have the weights X_M instead. The mismatch in the ESG part is therefore

$$
\begin{aligned}
\Pi - \Pi^* &= X_M^{'} L G_F - X_F^{*'} L G_F \\
&= \frac{\mu}{a\gamma_M}\Sigma^{-1} L G_F + \frac{c}{a^2\gamma_M} G_F^{*'} L^{'} \Sigma^{-1} L G_F + \frac{b}{a}\bar{X} L G_F \\
&\quad -\frac{\mu}{\gamma_F}\Sigma^{-1} L G_F - \frac{1}{\gamma_F^2} G_F^{'} L^{'} \Sigma^{-1} L G_F \\
&= \underbrace{\left(\frac{\mu}{a\gamma_M} - \frac{\mu}{\gamma_F}\right)\Sigma^{-1} L G_F}_{\Delta_1(\text{Return conflicts): fund manager only take } a \text{ shares}} \\
&\quad + \underbrace{\left(\frac{c}{a^2\gamma_M} - \frac{1}{\gamma_F^2}\right) G_F^{'} \Omega G_F}_{\Delta_2(\text{ESG interest conflicts): fund manager only take } c \text{ shares in pecuniary concerns}} \\
&\quad + \underbrace{\frac{c}{a^2\gamma_M}\left(G_F^{*'} - G_F^{'}\right)\Omega G_F}_{\Delta_3(\text{ESG assessment conflicts): fund manger follows social norm}} \\
&\quad + \underbrace{\frac{b}{a}\bar{X} L G_F}_{\Delta_4(\text{Benchmark benefit): fund manager should concern about the benchmark}}
\end{aligned}
$$

where $\Omega = L^{'}\Sigma^{-1}L$. Let's assume $\gamma_M = \gamma_F$ then we have $\Delta_1 < 0$ and $\Delta_4 > 0$. $\Delta_2 \propto (c - a^2)$, which is determined by the compensation structure of ESG fund manager. The only thing undefined is $\Delta_3 \propto (G_F^{*'} - G_F^{'})\Omega G_F$.

To estimate $\Pi - \Pi^*$, we shall estimate $G_F^{*'} - G_F^{'}$ in the 401(k) data. L,Σ,Ω, μ could be from the mutual fund database. a, b, c could be estimated from the fee regression. However, G_F could be heterogenous, like Pástor, Stambaugh and Taylor (2021) has mentioned. With different $G_{j,F}$, the $\Pi_j - \Pi_j^*$ would be different and serve to this paper.

Heterogeneous Investors with Two Types of Mutual Funds

- There are two types of funds: ESG fund (E) and index fund (I)
- Fund investors allocate the asset of $\Lambda = (\lambda_E, \lambda_I)$ fractions in E, and in N
- All information is under full commitment. We would relax this condition.

Mutual Fund Company Index fund managers face such optimization problems:

$$\max_{X_I} \mathbf{E}\left[-\exp(-\gamma_I w_I)\right] = \max_{X_I} \mathbf{E}\left[-\exp(-\gamma_I(aX_I'r - b\bar{X}r + \phi_I))\right] \tag{1.7}$$

where $a = \hat{a} + b$. The F.O.C gives that

$$X_I = \frac{1}{a\gamma_I}\Sigma^{-1}\mu + \frac{b}{a}\bar{X}. \tag{1.8}$$

ESG fund managers face the following optimization problem:

$$\max_{X_E} \mathbf{E}\left[-\exp(-\gamma_E w_E)\right] = \max_{X_E} \mathbf{E}\left[cX_E'L\bar{G} - dX_E'L\bar{G}\bar{G}'L'X_E + \phi_2\right] \tag{1.9}$$

The F.O.C gives that

$$X_E = \frac{c}{d}(L\bar{G}\bar{G}'L')^{-1}L\bar{G} \tag{1.10}$$

Investor Choice The investor i face the choice that

$$\begin{aligned}\max_{\lambda_{I,i},\lambda_{E,i}} \mathbf{E}\Bigg\{-\exp\Bigg[&-\gamma_i((\lambda_{I,i}(\underbrace{(1-a)X_I + b\bar{X}}_{\tilde{X}^*_{I,i}}) + \lambda_{E,i}X_E)'r - c\lambda_{E,i}X_E'L\bar{G} \\ &-\lambda_{I,i}\phi_1 - \lambda_{E,i}\phi_2) - (\lambda_{I,i}X_I + \lambda_{E,i}X_E)'LG_i\Bigg]\Bigg\}\end{aligned} \tag{1.11}$$

The F.O.C gives that

$$\begin{aligned}\tilde{X}_I'(\mu + \frac{1}{\gamma}LG_i) &= \gamma_i\tilde{X}_I'\Sigma(\lambda_{I,i}\tilde{X}_I + \lambda_{E,i}X_E) + \phi_1 &(1.12)\\ X_E'(\mu + \frac{1}{\gamma}LG_i - cL\bar{G}) &= \gamma_i\tilde{X}_E'\Sigma(\lambda_{I,i}\tilde{X}_I + \lambda_{E,i}X_E) + \phi_2 &(1.13)\\ &&(1.14)\end{aligned}$$

This give us

$$\begin{pmatrix}\lambda^*_{I,i}\\ \lambda^*_{E,i}\end{pmatrix} = \frac{1}{\gamma_i}\underbrace{\begin{pmatrix}\tilde{X}_I'\Sigma\tilde{X}_I & \tilde{X}_I'\Sigma X_E\\ \tilde{X}_I'\Sigma X_E & X_E'\Sigma\tilde{X}_E\end{pmatrix}^{-1}}_{M}\begin{pmatrix}\tilde{X}_I'(\mu + \frac{1}{\gamma}LG_i)\\ X_E'(\mu + \frac{1}{\gamma}LG_i - cL\bar{G})\end{pmatrix} \tag{1.15}$$

Market Clearing The market clearing gives up

$$\underbrace{\begin{pmatrix} X_I & X_E \end{pmatrix}}_{\boldsymbol{X}} \int_i \begin{pmatrix} \lambda^*_{I,i} \\ \lambda^*_{E,i} \end{pmatrix} di = \bar{X} \tag{1.16}$$

$$\frac{1}{\bar{\gamma}_i} \boldsymbol{X M} \begin{pmatrix} \tilde{X}'_I(\mu + \frac{1}{\gamma} L\bar{G}_i) \\ X'_E(\mu + \frac{1-c}{\gamma} L\bar{G}_i) \end{pmatrix} = \bar{X} \tag{1.17}$$

At the equilibrium, if we assume $\Delta P_i = LG_i - L\bar{G}$, we have

$$\frac{\partial \lambda_{E,i}}{\partial \Delta P_i} = \frac{1}{\gamma_i} \frac{AX_E - BX_I}{AC - B^2} \tag{1.18}$$

$$\frac{\partial U_i}{\partial \Delta P_i} = \frac{1}{2}(\lambda^*_{I,i} X^*_I + \lambda^*_{E,i} X^*_E) \tag{1.19}$$

where $A = \tilde{X}'_I \Sigma \tilde{X}_I$, $B = \tilde{X}'_I \Sigma X_E$, and $C = X'_E \Sigma \tilde{X}_E$. We have $AC - B^2 > 0$ given Cauchy Inequality. Let's assume $\bar{X} = 0$ and we have $\frac{\partial U_i}{\partial \Delta P_i} = \kappa - \Psi \Delta P_i$, where $\Psi = (X_I X'_E - X_E X'_I)\Sigma(X_I - X_E)$. U_i is maximized at $\Delta P_i = \Psi^{-1}\kappa$, and when there are more than one $\bar{G}$, under certain parameters we can find space $\Theta = (a, b, c, d)$ to make more investors better off. We would leave this part here given it is not the main contribution of this chapter and it helps guide future studies in this area.

Chapter 2

The Distribution Side of Insurance Markets

Li An, Wei Huang, Dong Lou, Jiaxing Tian, and Yongxiang Wang[1]

This paper studies the impact of sales channels on insurance product adoption. Specifically, we have access to novel policy-level data provided by one of the largest life insurers in China, where we observe detailed information on individual policy characteristics, investor characteristics, and sales channels. We exploit a regulatory change in 2014 that requires at least 20% of the contracts sold by bank insurance agents in each quarter to be qualified long-term insurance products. Exploiting a discontinuity-in-slope design, we show that bank agents falling below their target qualified ratios in the first two months of a quarter make up for the shortfall in the third month; conversely, bank agents that have exceeded their target ratios in the first two months do not alter their behaviors in the last month of the quarter. This shift in the qualified ratio in the last month of the quarter is entirely due to a product-composition change – switching from short-term unqualified life insurance products to long-term qualified annuity products. We further show that this switch is not achieved by changing the relative pricing of products or client compositions. Our results have useful implications for client welfare.

[1]We thank Ralph Koijen and seminar participants at Korea University Business School, Nottingham University, Renmin University, University of Bristol, University of Hong Kong, University of Reading, University of Texas Dallas for helpful comments. All remaining errors are our own.

2.1 Introduction

Households face a difficult search problem when choosing from a large menu of financial products – from bank savings products to mutual fund and insurance products, to more complicated structural products. To mitigate search costs, households often rely on the advice of intermediaries (e.g., local brokers and bank branches) in their financial decisions. These financial intermediaries serve a number of useful functions: to introduce a diverse range of financial products to otherwise uninformed clients, to explain the technical details of innovative, sometimes difficult-to-understand products, and to facilitate transactions.

An increasingly important concern with the interactions between households and their advisors is that financial intermediaries – in pursuit of own profits – may not act in the best interest of their clients. Prior research on sales and distribution channels in financial markets has focused mainly on the effect of incentive fees (e.g., commissions) paid to intermediaries on their clients' product adoption. Recent studies have shown, quite convincingly, that intermediaries and advisors recommend dominated products to their clients for higher sales commissions. Egan, Matvos and Seru (2019), for example, shows that between two products with identical payoffs, advisors often recommend the one with higher sales commissions (so lower net-of-fee returns) to households.

Financial intermediaries also have other business considerations than incentive fees (or commissions). For example, intermediaries often face tight or even binding regulatory and contractual constraints, which may impact their advice to clients. In this paper, we study one such regulatory constraint on sales of life insurance products in China. Unlike prior studies that examine nearly identical financial products with different sales commissions, our paper analyzes the extent to which the introduction of a regulation affects *the types* of financial products acquired by households (with different implications for client welfare).

To analyze this issue, we have obtained access to a unique dataset provided by one of the largest life insurers in China. Our data include a 10% random sample of all insurance policies sold in 9 large Chinese cities for the period 2012-2016. We observe the contractual documents which include policy details (product types, contract length, etc.), investor characteristics (age, gender, income, etc.) and sales channels (bank branches or personal agents). We then exploit a regulatory change in 2014 that imposes a binding constraint on the types of insurance contracts sold through banks. Specifically, before 2014, banks mostly sold short-term life insurance products; the 2014 regulation requires that at least 20% of the insurance products sold by each bank (aggregated at the province level) in each quarter must be long-term products. This new regulation provides a relatively clean discontinuity design to identify large swings in product adoption. A unique feature of our setting is that there is no change in the sales commissions paid to banks (or other distribution channels) across different insurance products around the policy shock, so the effect we document is orthogonal to the traditional channel of variation in sales commissions.

Our empirical strategy exploits branch-quarter variation in the distance-to-constraint and examines how bank branch sales near quarter-ends vary as a function of the distance-to-constraint, before vs. after the policy change. Note that the regulatory constraint is binding at the bank-quarter-province level; we argue that, in practice, banks in each city of the province have a specific target qualified ratio to accommodate the heterogeneity in clienteles across cities. We proxy for the city-level target ratio using the average qualified ratio across all branches within each city in the preceding year (which is then applied to all branches in the city for the next quarter).[2]

More specifically, our identification strategy exploits the fact that bank branches falling below their target qualified ratios in the first two months of a quarter have strong incentives (i.e., are likely required by the provincial headquarter) to make up for the shortfall in the third month; conversely, bank branches that have exceeded their target qualified ratios in the first two months have no incentive to alter their behavior in the final month of the quarter. In other words, we expect a discontinuous jump in the relation between the qualified ratio minus target ratio in the last month of a quarter and that in the previous two months as the lagged

[2] From our private conversation with our data provider, this is a common practice of how banks implement and achieve the regulatory sales target. Later we will provide empirical evidence to show that this construction is a good approximation of how banks distribute its provincial-level target to each of its branches.

ratio crosses from above to below the target ratio (i.e., from having slacks to having deficits).[3].

Our predictions are strongly borne out in the data. We start our analyses with new insurance contracts signed in the final month of each quarter (March, June, September, and December). For bank branches that fall below their target qualified ratios in the first two months of the quarter, the relation between the qualified ratio (minus the target ratio) in the final month and that in the first two months is statistically more negative than the same relation for bank branches above their target ratios in the first two months. The difference between the two is -0.528 with a t-statistic of -2.41. We also examine the same difference (between bank branches above and below the threshold) in the response coefficient for non-quarter-end months (i.e., the other eight months). Since the constraint is only binding at quarter ends, we expect to see a smaller difference. Indeed, the difference in the response coefficient in non-quarter-end months is statistically insignificant from zero, and with a wrong sign. We further separate all insurance contracts into two groups: life insurance contracts and annuity contracts, and find that virtually all our documented effect comes from changes in annuity products (with a difference in the response coefficient of -0.700 and a t-statistic of -4.17).

We then repeat our analyses with the total premium from all insurance contracts – both those signed in the final month of the quarter (new) and those signed before the final month (existing). We see very similar patterns. The response coefficient of the distance-to-target-ratio in the final month of the quarter on the lagged distance for bank branches above the threshold is statistically more negative than that for bank branches below the threshold. Again, there is no discernible difference in the response coefficient for non-quarter-end months.[4]

In a placebo test, rather than using the average qualified ratio across all branches in the city from the preceding year, we add a random noise to the average ratio as the new target ratio. Our pattern gets monotonically weaker as we increase

[3]In an ideal setting, we expect the slope to be exactly 0 for bank branches with qualified-ratio slacks in the first two months and a slope of -1 for branches with qualified ratio deficits (on a value-weighted basis)

[4]We also provide evidence that bank branches with qualified-ratio deficits in the first two months of a quarter delay the payment of their existing short-term insurance products to the following quarter to improve the qualified ratio in the current quarter.

the variance of the noise term. This provides support for the premise of empirical design that bank branches in the same city have similar target qualified ratios.

An interesting feature of the insurance contracts offered by our firm is that while most life insurance contracts have a short maturity (less than 10 years), so do not qualify for long-term investments, virtually all annuity products have a maturity longer than 10 years, so qualify as long-term investments. We find that the shift in the qualified ratio almost driven entirely by a composition change – to switch from short-term unqualified life insurance products to long-term qualified annuity products, with the total premium unchanged.

An obvious concern with our results so far is that households may take out long-term annuity products but then decide to cancel these contracts shortly after, so there is no change in the effective investment horizon. At least in our sample period, from 2012 to 2016, we do not see an increase in the lapsation rate for long-term contracts signed in the final month of the quarter in response to a target-qualified ratio deficit. We further show that the switch from life insurance products to annuity contracts introduces changes in mortality delta. Although this is mostly mechanical (as annuity contracts have larger life components than life insurance products, the change in mortality delta have important welfare implications.

Although our analysis focuses on a particular regulatory change in the Chinese insurance market, our results have broader implications for other financial products in other countries. Sales team/distribution channels often face quotas and sales targets (for example, a sales person needs to sell $X of a product by year end); these sales targets can have very similar impacts on sales and production adoptions as the regulatory constraint examined in this paper.

A natural follow-up question is how do distribution channels achieve their sales targets? First, distribution channels, together with product providers, can change the (relative) pricing of products. Second, distribution channels can spend more effort/resources courting clients more suited for the product in question, therefore forgoing other types of clients (so a change in client composition). Third, distribution channels can persuade their clients to adopt certain products (not through pricing, but through communication and perhaps manipulation). This is also a substitution away from other products, but with no change in client characteristics.

To differentiate the first channel from the other two channels, we examine the pricing markups of different products. Following the standard procedure to calculate insurance products' markups in prior research, we find no significant change in markups of insurance product pricing in the last month of the quarter in a way that would lead investors to switch from short-term insurance products to long-term products.

To differentiate the second channel from the third, we examine changes in client characteristics that are associated with insurance purchase decisions. Our firm collects three client characteristics: gender, age, and income.[5] We find no significant change in client characteristics near quarter ends as a function of the qualified ratio in the first two months of the quarter. In other words, while there is a change in the type of products sold, at least based on important observables (which our firm cares about), there is no discernible change in investor characteristics or composition.

In our last set of analyses, we repeat our empirical exercises on the subset of contracts sold through the personal agent (PA) channel. We do not find similar patterns in qualified vs. non-qualified insurance premium in the last month of the quarter. This is consistent with the fact that personal agents are not subject to the new regulation.

In sum, our analysis (with the difference-in-slope test) reveals that as bank branches fall below their target qualified ratios in the first two months of the quarter, they increase the sales of qualified long-term products, which are predominantly annuity products, and reduce the sales of unqualified short-term products, which are mainly life-insurance products. We also provide suggestive evidence that bank branches achieve this switch in composition via persuasion, rather than by changing the relative pricing of their products or changing the client compositions.

The main contribution of our paper is twofold. First, prior research on the incentives of the distribution channel shows that distributors persuade their clients to

[5]The firm also collects information on clients' marital status and the number of children, but this information is missing for about half of the policyholders.

adopt "overpriced" products (compared to other products with identical or similar payoffs), because these overpriced products pay a higher sales commission. Our paper instead shows that distributors also promote "wrong" products. In particular, we show in a discontinuity setting that nearly identical clients buy short-term life insurance in the pre-2014 period but long-term annuities in the post-2014 period. Almost by definition, short-term life products are vastly different from long-term annuities, both in terms of contract duration and mortality benefits. Our results therefore highlight the fact that many retail investors may have very limited financial knowledge about *what kind of financial products* are suitable for their goals and preferences. Second, while prior studies examine the relation between incentive/commission fees and products recommended, our paper focuses on regulatory and contractual constraints imposed on financial intermediaries.

Related Literature Our study contributes to the body of literature that underscores the importance of supply side factors in insurance markets. Koijen and Yogo (2015) show that insurance companies sold their policies at discounted prices when they suffered from balance sheet shocks. This heavy price discount on one hand attracts demands from policy buyers and on the other hand increases accounting profits (hence ratings) as long as it is still above the reserve value set by statutory reserve regulation in the United States. Ge (2022) extends Chevalier's pioneering study on how financial constraints affect product pricing in the context of insurance market and also Koijen and Yogo (2015). This study shows that premiums fall (rise) for life policies that immediately increase (decrease) insurers' financial resources. In a similar vein, Ge and Weisbach (2021) study how operating losses (e.g., due to adverse weather conditions) for insurers affect their asset portfolio allocations. Koijen, Van Nieuwerburgh and Yogo (2016) show that life insurance companies shifted their liabilities to shadow reinsurers which are less regulated and unrated off-balance-sheet entities within the same insurance group. While such shifting reduced the marginal cost of issuing policies, it increases the default risk of the industry.

2.2 Data and Empirical Strategy

2.2.1 Data and Summary Statistics

We exploit a unique dataset provided by one of the largest life insurers in China. The dataset contains a 10% random sample of all contracts signed in the period of 2009-2016 in 9 large cities in China: Beijing, Shanghai, Guangzhou, Chengdu, Nanjing, Wuhan, Shenyang, Zhengzhou, and Lanzhou. The total insurance premium in our sample was over 3B RMB in 2016. For our main analyses, we focus on the period from 2012 to 2016, centering on the policy change in 2014.[6]

The dataset contains comprehensive details from the contractual documents for each policy, encompassing several sets of information. Firstly, it provides insights into policy details, such as product type, contract length, annual premium and premium payment period. In total, there are 231 products included, categorized into four product types: life insurance, annuities, health insurance, and accident insurance. Secondly, the dataset encompasses investor characteristics, including gender, age, annual income, occupation, and marital status. Thirdly, we can track all transactions related to the contracts, such as premium payments, contract lapsation, and insurance claims.

Important for our purpose, we also observe information on the distribution channel. Insurance policies are distributed through diverse channels. Figure 2.1 shows the premium revenue in our sample from each distribution channel. The majority of observations, over 95%, are sold by either banks or personal agents. We observe sales by each bank branch or personal agent with anonymous IDs, but we do not have information on the bank identity.

Table 1 reports summary statistics and describes the basic facts of our sample. Panel A reports the time-series insurance sales in our sample. We focus on the number of contracts and the premium revenue from the new contracts. All insurance products are classified into four categories: life insurance, annuity, health

[6]The total premium collected in each quarter comprises premiums from newly issued policies as well as those from existing policies. By 2012, the impact of policies issued before 2009 (not in our sample) diminishes significantly.

insurance, and accident insurance. In our sample, there is no universal life products or whole life products. Even the term-life products would pay back the insurance value at maturity. Given all products or contracts combine life insurance and annuity components in our sample, if the maturity exceeds 10 years and the expected present value of all annuity claims is not less than the insurance value, they are categorized as annuities. The remaining combined products with different characteristics are attributed to the life insurance category. Any life insurance or annuity products that also include health and accident components are classified as life insurance or annuity only.

The majority of premium contributions are attributed to two product types: life insurance and annuities, whereas accident and health products (A&H) constitute less than 4% of the total premium across all years in our whole sample. Life insurances take more than 80% of the premium while this number declines after 2014. Annuities stay below 16% before 2014, and surge after that. In terms of the number of contracts, this trend is similar as the premium revenue.[7] Accident and health insurance are higher in the number of contracts while they are usually in lower value.

Shifting the focus only on the bank as the sales channel. Panel B reports the summary of insurance sales of each branch in each month. In average, a branch would sell 0.86 number contracts per month in our sample, and given it is a 10% sampling, this would be 8.6 new contracts in reality. These new contracts contribute more than 83.78 thousand RMB while the lapsation rate within the next 12 months after the new contract signing would be lower than 1%. The duration of life insurance is quite shorter than the annuities (7.41 v.s. 20.65), representing that life insurances are usually short-term products.

For these new contracts sold, Panel C reports the contract and buyer characteristics. The markup for each contract is based on the ratio of the expected claim amount within the contract maturity to the expected present value of the total

[7]Given the life insurances are usually in high value in our sample, the proportion of life insurances takes around 50-60% before 2014 and below 35% after 2014.

premium:

$$M_{i,t} = \left\{ \sum_{s=t+1}^{t+T} \left[A_{i,s}(1-\pi^*_{D,s|i}) + C_{i,s}\pi^*_{D,s|i} \right] \pi_{L,s-1|i,t} d_{s|t} \right\} / \left(\sum_{k=t}^{t+T_0} P_{i,k}\pi_{L,k|i,t} d_{k|t} \right). \tag{2.1}$$

$\pi_{L,k|i,t}$ is the probability of being alive at period k conditional on the insuree's characteristics of contract i and being alive at t. $\pi^*_{D,s|i}$ is the probability of death at period s conditional on being alive at $s-1$ for the insuree of contract i. These mortality rates are based on the Sixth National Population Census of China in 2010 and are heterogeneous in insuree's age and gender. $A_{i,s}$ is the expected annuity payment at period s, including annuity itself and dividend if contracted. $C_{i,s}$ is the death claim for contract i at period s. $P_{i,k}$ is the premium amount at period k. These numbers, $A_{i,s}$, $C_{i,s}$, and $P_{i,k}$, are hand-collected for each specific product and vary by insurance value and premium choices. T and T_0 stand for contract maturity and premium payment period respectively. If it is a whole life insurance, T is set to make the insuree's age 110 at that time.[8] $d_{k|t}$ is the discount rate based on the $(k-t)$-year treasury bond rate, namely China Government Bond yield, at t. The data is from the Wind financial database. For any rate $k-t$ without a corresponding treasury bond, we linearly interpolate from the two bonds with the closest maturities.

This markup has a mean value of 1.085 across all new contracts and it is quite concentrated within the range of $(0.94, 1.22)$. To capture whether the contracts are at fair prices, we also run a cross-sectional regression of these markups on some contract and insured characteristics and age a zero mean residual. This will be discussed in detail in Section 2.4.2. The buyers are mainly females, and the ages are just below 49. The annual income of these buyers is 70k RMB but is highly volatile.

To evaluate whether the buyer is investing in short-term dearth risk or long-term aliveness gain, we also calculate the mortality delta (δ), following Koijen, Van Nieuwerburgh and Yogo (2016). The δ is defined as the payoff that a policy

[8]Most contract policies in our sample set the maximum duration of insured age to be 110. In fact, the terms in the numerator after $t > 90$ would decrease dramatically to zero because the conditional mortality rate $\pi^*_{D,s|i}$ would be close to 1.

delivers at death relative to being alive in the next period:

$$\delta_{i,t} = C_{i,t+1} - \left\{ A_{i,t+1} + \sum_{s=t+2}^{t+T} \left[A_{i,s}(1-\pi^{*}_{D,s|i}) + C_{i,s}\pi^{*}_{D,s|i} \right] \pi_{L,s-1|i,t+1} \frac{d_{s|t}}{d_{t+1|t}} \right\}. \tag{2.2}$$

By normalizing with the insurance value, the measure Delta/Value have a mean of 0.18 but could have a variation range from 0.04 to 0.56.

2.2.2 Identification Strategy

In 2014, the China Banking Regulatory Commission and the China Insurance Regulatory Commission jointly issued a notice to standardize the practices of insurance sales in the bank agent channel (CBRC [2014] No.3). Before 2014, banks primarily sold short-term life insurance products; the new regulation, implemented in April 2014, mandates that bank-insurance sales channels to prioritize long-term products. Specifically, each bank is required to ensure that at least 20% of the quarterly premiums (aggregated at the province level) originated from 'qualified' long-term insurance products. These qualified products include annuities, long-term life insurance, health insurance, and accident insurance, all with terms exceeding 10 years. Figure 2.2 shows that all of the annuities in our sample are qualified and around 8% of life insurances are qualified.

Our identification strategy essentially exploits the branch-quarter level variation of compliance pressure and examines how bank branch sales respond to the distance-to-target before and after the policy change. The regulatory constraint is binding at the aggregated bank-quarter-province level. We argue that, in practice, each bank in each city has a specific target qualified ratio to accommodate the diverse clientele it may encounter. This target ratio is proxied by the average qualified ratio across all branches within the city in the preceding year, and applies to all branches for the next quarter.[9] If a branch is below this target in the first two months of a quarter, it has a strong incentive to make up for the shortfall in the

[9]From our private conversation with our data provider, this is a common practice how banks implement and achieve the regulatory sales target. Later we will provide empirical evidence to show that this construction is a good approximation of how banks distribute its provincial-level target to each of its branches.

third month; conversely, if the branch surpasses this target, there is no incentive to alter its behavior.

Specifically, we employ the following regression model:

$$Y_{i,j,t} = \beta_1 D_{QR^{L2}_{i,j,t}<C_{j,t}} \times (QR^{L2}_{i,j,t} - C_{j,t}) + \beta_2 D_{2014} \times D_{QR^{L2}_{i,j,t}<C_{j,t}} \times (QR^{L2}_{i,j,t} - C_{j,t}) + \sum_{y=2012}^{2016} \gamma^1_y \times D_y \times (QR^{L2}_{i,j,t} - C_{j,t}) + \sum_{y=2012}^{2016} \gamma^2_y \times D_y \times D_{QR^{L2}_{i,j,t}<C_{j,t}} + \theta_t + \eta_i + \epsilon_{i,t} \tag{2.3}$$

$Y_{i,j,t}$ is our outcome variable of interest and is the qualified ratio (premium of qualified contracts/premium of all contracts) in the third month of the quarter by branch i in bank-city combination j in quarter t in most analyses. $C_{j,t}$ is the bank-city level target, proxied by the average qualified ratio in the previous four quarters across all branches under that bank in that city. $QR^{L2}_{i,j,t}$ is branch i's qualified ratio in the previous two months (the first two months in the quarter). $D_{QR^{L2}_{i,j,t}<C_{j,t}}$ is a dummy variable indicating whether branch i's qualified ratio in the first two month falls below the threshold. D_{2014} is a dummy variable that is equal to 1 after 2014 Q2, indicating time periods before and after the policy change.

Our main variable of interest is the triple interaction term $D_{2014} \times D_{QR^{L2}_{i,j,t}<C_{j,t}} \times (QR^{L2}_{i,j,t} - C_{j,t})$. We conjecture that the distance to target in the first two months, $QR^{L2}_{i,j,t} - C_{j,t}$, would have an negative impact on the sales composition (qualified ratio) in the third month, only when it's after the policy change and if the branch's qualified ratio in the first two month falls below the threshold ($\beta_2 < 0$). $D_{QR^{L2}_{i,j,t}<C_{j,t}} \times (QR^{L2}_{i,j,t} - C_{j,t})$ is the double interaction term and β_1 captures the effect of distance to target on the third-month sales when a bank branch falls short of the threshold before the policy change. We expect β_1 to be zero. Instead of including the other two double-interaction terms, $D_{2014} \times (QR^{L2}_{i,j,t} - C_{j,t})$ and $D_{2014} \times D_{QR^{L2}_{i,j,t}<C_{j,t}}$, we control for the interaction terms between $(QR^{L2}_{i,j,t} - C_{j,t})$, $D_{QR^{L2}_{i,j,t}<C_{j,t}}$ and a series of year dummies; this serves as a finer version of control that allows the effect to change every year. Additionally, we include quarter fixed effects θ_t and bank branch fixed effects η_i.

2.3 Main Results

2.3.1 Distance-to-Constraint and Qualified Ratio Near Quarter-Ends

We start by analyzing how branch-quarter level compliance pressure affects the sales composition of long-term versus short-term products in the last month of each quarter. Table 2.2 reports estimation of β_1 and β_2 in Eq. 2.3 using the qualified ratio of newly signed contracts as the dependent variable. We focus on premiums from new contracts as they are more easily manipulated by sales force compared to revenues from existing contracts.

In Panel A, the dependent variable is the abnormal qualified ratio from new contracts. Specifically, it is calculated as the qualified ratio (qualified premium/total premium) of newly signed contracts across all product types for each branch in the last month of each quarter, minus the same ratio from the previous four quarters for the same branch. By subtracting the bank branch's own historical ratio, we control for time-varying branch-specific factors that may drive sales composition. Following the regulation reform, a negative deviation from the target leads to a significantly higher qualified premium ratio ($\beta_2 < 0$) in the last month of each quarter, as indicated in the first column. This finding aligns with our conjecture, and we also observe an insignificant β_1, suggesting that this pattern did not exist before 2014. These results indicate that a bank branch's distance-to-target in the initial two months of a quarter influences its sales composition in the subsequent third month.

Economically, when a branch's qualified ratio is lower than the bank-city target of 1% in the initial two months, they increase the ratio by 0.528% in the last month. Though lower than one, $C_{j,t}$ is a proxy, with the constraint binding at the bank-city level. At the same time β_2 captures the response of branches below the target compared to those above, suggesting a $4.876 (= 83.78 * 0.582\% * 10)$ thousand RMB increase per 1% below the target.

To conduct a placebo test, we randomize the sequence of months within a quarter. We examine sales composition in a non-quarter-ending month (e.g., January, we call them as placebo months) and compute the distance-to-target based on the remaining two months within the same quarter (e.g., February and March). Compared with the last month of each quarter, the second column of Panel A indicates that branches with negative distance-to-target have an insignificant β_2 in the placebo months. Compared with the coefficients of β_2 in the sample month and placebo month, the last column shows that the response is significantly different. Given that the constraints bind at the end of each quarter, these results not only support our conjecture that branches respond to the constraints at quarter-ends, but also validate our assumption regarding the proxy of the target ratio.

The question here is which type of insurance products are sold due to the binding constraint. In Panels B and C, we use the qualified premium from life insurance and annuities, respectively, to divide the total new premium in a given month as $y_{i,t}$, also subtracting these ratios from the previous four quarters. In Panel B, we find that the qualified premium from life insurance doesn't change in response to the binding constraint, supported by an insignificant β_2 in the sample month. In Panel C, the significant β_2 indicates that the proportion of premiums from annuities, all of which are qualified, increases when the branch falls behind the target ratio. The insignificant β_2 in the placebo month and the significant difference between the sample and placebo months indicate that branches tend to sell more annuities to catch up with the target ratio of the binding constraint. Combining this with the previous result in the life insurance, branches are more likely to substitute unqualified life products with annuities while keeping the sales of qualified life products unchanged.

The reason for such substitution may stem from sticky menu costs. In Panel B of Table 2.1, the average (median) duration of life products banks sold is 7.41 (6) years, indicating that, on average, life insurances are unqualified. Although this is an ex-post summary, one possible explanation is that the pool of life products banks could sell is predominantly unqualified. Even after the regulation reform, it may be costly to change the menu within the two-year period of our sample. Additionally, annuities have similar cash flow durations to short-term life products compared to long-term life products[10]. If branches aim to sell more qualified products due to the binding constraint through persuasion, it may be easier to sell annuities rather than long-term life products. Given our product definitions, most annuities include a life component as well, covering insured death during the policy period.

The response not only exists in new premiums but also in qualified ratios in overall premiums, as reported in Table 2.3. In Panel A, $y_{i,t}$ represents the qualified premium ratio from all contracts, subtracting this ratio from the previous four quarters to capture qualified revenue from both new and old contracts. β_2 remains significantly negative in sample months and close to zero in placebo months, with a significant difference suggesting consistent responses. The magnitude of β_2 is almost identical to Table 2.2 because most premium manipulation comes from new product sales. We will examine premium from old contracts in the following test. β_1s are still insignificant, supporting that responses are only due to the regulation reform.

In Panels B and C, we conduct similar regressions as in Table 2.2. The significantly negative β_2 in the first column of Panel B suggests that even qualified premium from life insurance increases to compensate for the shortfall in the third month, though differences between sample and placebo months are insignificant. Results in the annuity part resemble those in new contract premiums. Branches receive more qualified premiums from both life insurance and annuities, slightly different from new contract premiums, possibly because branches also lower premiums from existing unqualified life products by not urging buyer payments.[11]

Given that the 2014 regulation necessitates compliance at the aggregated bank-province level, there arises a valid concern regarding whether our construction of sales target at the bank-branch level accurately reflects how banks distribute and execute this regulatory requirement among their branches. To address this concern, we empirically investigate the validity of our construction of the target.

[10] Looking at product durations, annuities have a median duration of 12.5 years, while long-term life products have a duration of 25.8 years.

[11] We also show that, the main responses to the regulation reform at the quarter ends exist in more developed cities in China. The results are shown in out appendix in Table 2.A.1 and Table 2.A.2. One possible reason for that is that the city-bank level targets are closer to the regulation requirement given their large premium revenue. In this way, our proxies are more accurate. And the purchasing ability in these cities are higher than less developed ones.

Specifically, we add a random component with different standard deviations to the target each bank branch faces in each quarter and repeat our analyses. The idea is that the correct target should produce the largest kink in slopes; if the target we construct reflects the true threshold, the results should become weaker as we add more noise to the calculation.

Table 4 presents regression estimations of βs using random thresholds. Given the target ratio should not deviate far from 20%, which is the explicit requirement of the regulation reform, we generate random thresholds using the process:

$$C^*_{i,t} = 0.2 + \sigma\epsilon^{'}_{i,t}.$$

Here, $\epsilon^{'}_{i,t}$ is simulated from the standard normal distribution $N(0, 1)$. σ is set to be 0.1, 0.2, and 0.3 respectively in Panels A, B, and C. β_2 declines with respect to σ, and its significance also declines. When σ is lower, the simulated thresholds are closer to the explicit requirements and our proxies. Hence, finding marginally significant negative β_2 estimations in Panel A is unsurprising. However, when the cutoffs deviate far from the requirement and randomly from our proxies, the kink should disappear. Conversely, the correct threshold should have the largest kink in slopes, providing a necessary condition for our assumptions.

We then investigate whether branches, facing binding constraints, delay unqualified premiums from existing contracts. We find the qualified ratio from existing life insurances also increases in Table 2.3. Table 2.5 uses the delayed unqualified ratio as the dependent variable $y_{i,t}$. The numerator uses the scheduled but delayed premium from unqualified existing premiums. Even short-term life insurances, for example 5-year ones, could allow multiperiod premium payment, e.g., in three years. The denominator includes the total scheduled unqualified premium, including both received and unreceived amounts in the given month. We also subtract this ratio from the previous four quarters in $y_{i,t}$. β_2 is negatively significant. Economically, with 1% lower than the target, the branch tends to delay 0.44% of their unqualified scheduled premium to the following months and quarters. However, the difference of β_2 between the sample month and placebo month is marginally insignificant, mainly because it is challenging to adjust the target ratio through this channel. Investors are difficult to persuade if they have already scheduled the premium, and such payments are sometimes automatic through bank account transfer.

2.3.2 Insurance Substitution and Contract Lapsation

As we document that bank branches increase their qualified premium ratio when they fall behind the target after the policy change, a natural question is whether they sell more qualified contracts solely or substituting the qualified contracts for the unqualified ones. This is important because the two channels lead to different implications on the supply and demand sides. We show that it is more likely that banks simply substitute the unqualified ones with qualified contracts, without changing the bank's total premium revenue in Table 2.6. Such substitution would not lead to a surge in lapsation rate of new contracts sold.

Table 2.6 uses the logarithm of the absolute value of premium revenue in each branch as $y_{i,t}$. In the first panel, the insignificant β_2 indicates that binding constraints are less likely to change the total revenue and there are no differences in both sample months and placebo months. Panels B and C support this substitution mechanism. In Panel B, we focus on the premium from life insurances and find a significantly positive β_2 in sample months. This suggests that the binding constraint leads branches to sell fewer life contracts, given part of the life products are unqualified. This, as shown in Panel C, has been compensated with more sales in annuities. The negative β_2 in annuity premium regression shows that branches make up the qualified ratio by more premium from annuities, which are all qualified.[12]

The substitution, on the one hand, changes the pool of products the branch offers or recommends. On the other hand, such responses would affect buyer decisions or target a different pool of buyers. We would discuss this in the following Section 2.4. Before moving on and without distinguishing any potential mechanism, we

[12] We show that given such substitution, the mortality delta is decreasing for the binding branches after the regulation reform, at the end of each quarter. Table 2.A.3 on our appendix shows that, cosmetically, each branch sells more annuities making on average buyers invest more in the insured aliveness than the dearth.

first check whether the lapsation probability of these contracts changes. Table 2.7 uses the lapsation rates as the dependent variables. Panel A uses the equal-weighted lapsation rates, i.e., the number of lapsed new contracts over all new contracts, as the dependent variable. We define the lapsed new contract as the contract policy signed in this given month but lapsed for the buyer's personal reasons within the next 12 months, ex-post. All estimators are close to zero and insignificant, suggesting that although the binding constraint changes the insurance sale structure, the total lapsation rates are unlikely to change. This result is robust if we focus on the value-weighted lapsation rates, i.e., the total premium of lapsed new contracts over the value of all new contracts, in Panel B.

2.4 Mechanisms: How Do Banks Achieve Sales Targets?

In this section, we investigate the mechanisms that help us understand how distribution channels achieve their sales targets. There are several possibilities. Firstly, distribution channels, in collaboration with product providers, may alter the pricing of different products, thereby influencing the relative attractiveness to investors. Secondly, these channels might direct their efforts towards attracting clients more inclined or suitable towards specific products, resulting in a different composition of policyholders. Thirdly, distribution channels could persuade their clientele to adopt certain products through communication, possibly manipulation. This could lead to a shift away from other products, even without any alteration in client characteristics. We examine each of the possible mechanisms in the following analyses.

2.4.1 Do Banks Change Relative Pricing?

Banks could adjust the relative price of qualified contracts to attract buyers, evaluated as the markup described in Eq 2.1. We calculate the $Markup_{i,k,t}$ of each contract and attempt to explain such relative price using policy and insured characteristics. Policy characteristics include whether the policy pays dividends to the insured and the contract duration. In China, both life insurance and annuities may pay dividends based on the insurance company's investment profit, usually variable but claimed when signing the new contract. Insured characteristics include gender, age, and annual income. Pricing functions vary across product types and time periods.

In Table 2.8, we use linear regression to approximate the pricing function, separating observations into life insurance and annuities. Regressions are divided into two sample periods: before 2014 (pre-policy period) and after 2014 (post-policy period). Interaction terms of policy and insured characteristics are included. Annual income negatively impacts annuity markup, while insured age negatively affects life insurance markup. Male insured individuals tend to have higher markups due to higher death risk. Markups for both products increase with duration, indicating higher expected policy claims probability over time. If a life insurance product pays dividends, markups are higher for low-income, male, and older insured individuals, especially after 2014, likely due to longer-term products. Conversely, annuity markup's numerator is negatively correlated with the mortality rate. Estimators for dividend-related terms are complex, as dividend payment is negatively correlated with the denominator in Eq 2.1 but may be positively or negatively correlated with the numerator through the mortality rate or policy payment.

Based on valid characteristics explaining markup, we run explanatory regressions yearly and separately for life insurance and annuities to obtain regression residuals. These residuals, at least partially, reflect the pricing freedom negotiable between bank distributors and buyers. We focus on both markups and the residuals as the relative price of insurance.

In Table 9, we use markups and their residuals as the dependent variables in our regression framework, with contract-branch-quarter level observations. Although dependent variables are at the contract level, independent variables remain at the branch-quarter level. This allows us to assess whether the regulation reform changes the price of each contract. In the first three columns of each panel, point estimates of β_1 and β_2 in both sample and placebo months are insignificant and close to zero. This suggests that markups across different branches during pre- or post-regulation periods are stable and unrelated to the binding constraint. To check if this is robust for qualified and unqualified products, we separate contractual observations into two groups and run regressions in these subsamples. The following six columns still show no significant β_1 and β_2, indicating branches are likely not lowering prices of qualified products or increasing prices of unqualified ones.

2.4.2 Shift of Clientele or Persuasion Aimed at the Same Client Base?

Given that branches are not changing their total premium revenue and not adjusting the relative prices of single products, the question arises: who is on the buyer side? Branches may change the pool of buyers to meet the target ratio, for example, shifting potential buyers from old to young, who are more likely to buy annuities. We use buyer characteristics as $y_{i,t}$s at contract-level observations in our empirical framework. In Panel A of Table 2.10, we run a Probit regression model and find almost all β_2s are insignificant in either placebo or sample months. This suggests that after the regulation reform, branches with binding constraints are not selling more to males or females compared to average cases, in both qualified and unqualified products. The only outlier is unqualified cases in the placebo month, but the significance is marginal.

Using buyer age and annual income yields similar results in Panels B and C. While it's not sufficient evidence that the potential client pool remains unchanged, buyer characteristics remain the same for branches with and without target ratio pressure, regardless of before or after the regulation reform. In conclusion, at least the pool of clients remains unshifted. One possible way to sell more qualified products is through persuasion to convince the same client to choose the qualified product.

2.5 Additional Results: The Personal Agent Channel

In this section, we validate our empirical findings by testing the impact of counterfactual target ratios on personal agents' insurance sales. We have provided several pieces of evidence supporting our basic assumption that branches would catch up to the target ratio at the end of each quarter after the regulation reform. However, there are alternative assumptions or endogenous reasons that could lead to similar findings. For example, it could be the time-series trend of product selling, a pure supply-side reason, or clients' general preferences at the end of each quarter, a pure demand-side reason. If it is not a distribution-side change, this would challenge our findings. We shift our focus to the personal agent sales channel to conduct similar but placebo analyses to demonstrate that our findings only apply to the bank sales channel.

We adopt an identical empirical strategy but utilize personal agent salesman-level observations. Each salesman, corresponding to one institution agent in our sample, is equivalent to a bank branch. The counterfactual target qualified ratio is assumed to bind at each institution agent in each city, still at the end of each quarter. Following the method in the bank channel, we aggregate small institution agents with previous revenue less than 100,000 RMB in the previous year into a "virtual" agent in each city and each quarter. The target ratios are proxied with the average qualified ratio in the previous four quarters across all salesmen under that institution agent in that city. In Panel A of Table 2.11, we conduct the identical study as in Panel A of Table 2.2, with the same dependent variable construction. The lack of significance of β_1 and β_2 in both sample and placebo months indicates that our target ratio does not have a significant impact on the personal agent's sales propensity. To determine whether this holds true for both life insurance and annuity sales, we conduct regressions identical to Panel B and C of Table 2.2 and report the results in Panel B and C of Table 2.11. Most of the β_2 estimates in this regression are insignificant and close to zero, particularly in the sample month column, reinforcing the validity of our findings in the bank channel.

To assess the response of premium and lapsation rates to the "virtual" target ratio, we follow the methodology of Panel A of Table 2.6 and Panel B of Table 2.7 in the personal agent channel observations. Given that this target ratio is counterfactual, we expect it to have no impact on salesmen's total premium and contract lapsation rates. The results in Table 2.12 support this claim. In Panel A, $y_{i,t}$ represents the total premium revenue of each salesman in a given month. In Panel B, the dependent variable is the lapsation value over the total value of the new contracts sold in this given month. Almost all β_2 estimates are insignificant, and those marginally significant are close to zero or lack economic significance. The constraint, if valid in the bank channel, has no impact on any outcomes in the personal agent channel. This could be served as another necessary evidence of the causal impact of regulation reform on the bank distributor channel.

2.6 Conclusion

This paper studies the impact of distribution channels on insurance product adoption. We exploit a regulatory change in 2014 that requires at least 20% of the insurance products sold by bank insurance agents in each quarter to be qualified long-term insurance products. Leveraging an unique policy-level data provided by one of the largest life insurers in China, we analyse how branch-quarter variation in distance-to-constraint affect sales composition in each of the bank branches. Exploiting a discontinuity-in-slope empirical design, we show that bank agents falling below their target qualified ratios in the first two months of a quarter make up for the shortfall in the third month; conversely, bank branches that have exceeded their target ratios in the first two months do not alter their behavior in the final month of the quarter. This shift in the qualified ratio in the last month of the quarter is entirely due to a composition change – to switch from short-term unqualified life insurance products to long-term qualified annuity products.

Further analyses on the mechanisms show that this change in sales composition is achieved mainly via persuasion, rather than by changing the relative pricing of the products or changing client compositions. Our results highlight the fact that many retail investors may have very limited financial knowledge about what kind of financial products are suitable for their goals and preferences. Constraints that distribution channels face can therefore generate a large impact on financial product adoption and investor welfare.

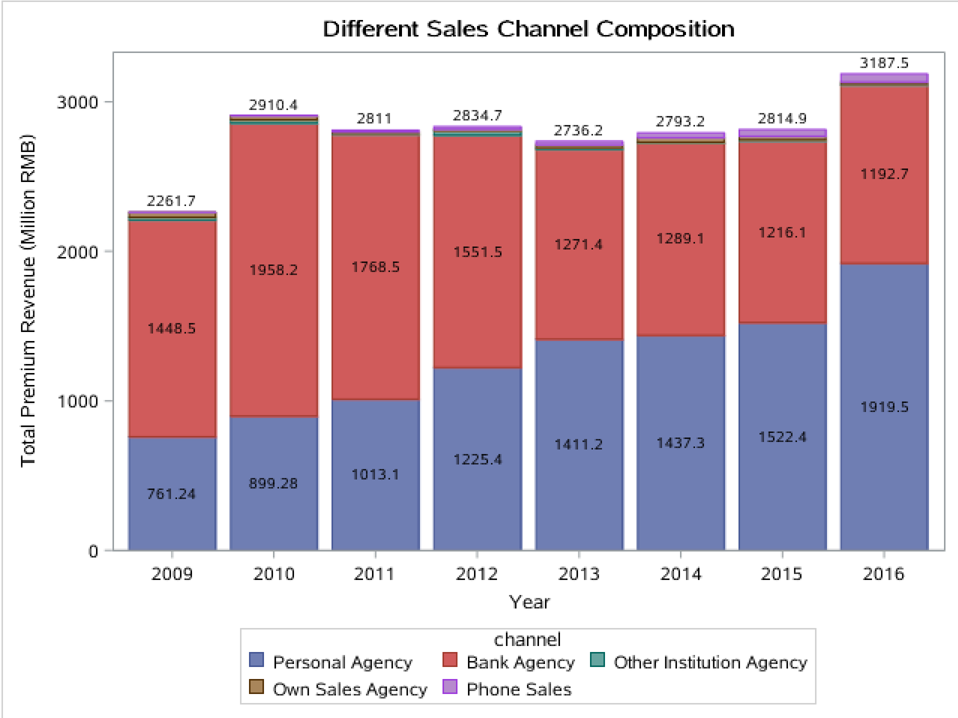


Figure 2.1: Premium revenue from different distribution channels

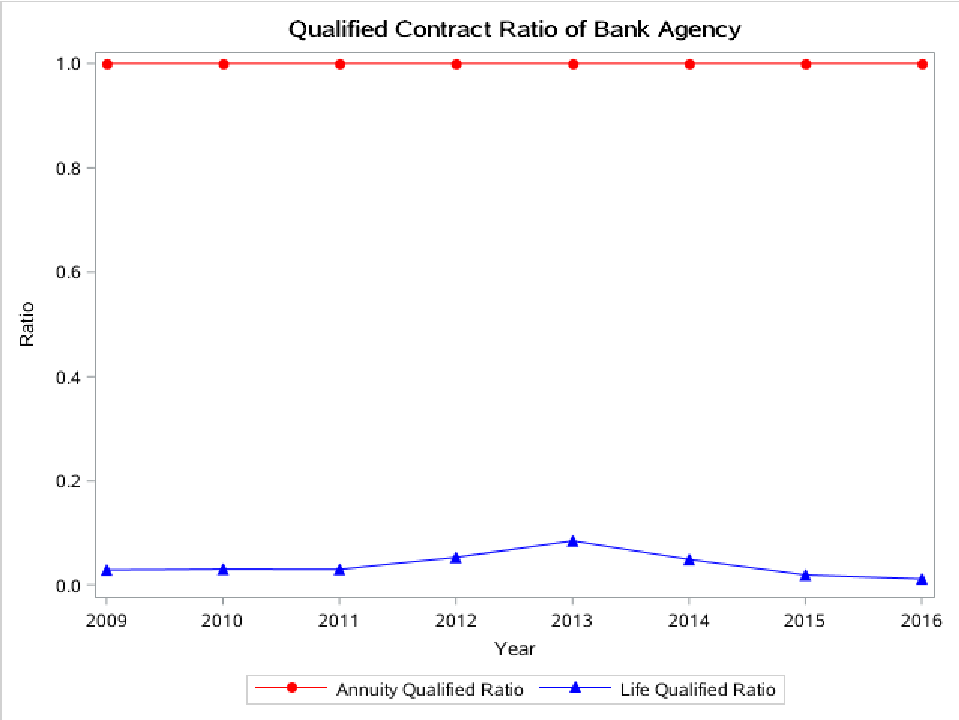


Figure 2.2: Qualified ratio of new contracts sold by bank channel

Table 2.1: Summary statistics

This table reports the summary statistics of insurance selling of our sample for the period 2009 to 2016.

Panel A reports time-series insurance sales by SalesmanIDs from all sales channels. In the second column, the total number of insurance contracts sold in that given year is shown, along with the proportion of each type of insurance product. The term "A&H" stands for accident and health insurance combined. Column 6 reports the total insurance premium in million RMB from these new contracts in the given year. For contracts with multi-period payments, only the premium occurring in that year is taken into account. The following three columns show the proportion of each type of insurance product contributing to the premium.

Panel B reports premium revenue, average lapsation rate, and contract duration of each bank branch for each month. Qualified premiums refer to those from long-term products required by the 2014 regulation reform. In the second part of the panel, only branches with new contracts sold in a given month are reported. Lapsation refers to the proportion of new contracts lapsed within 12 months due to personal reasons. The last part of the panel focuses on branches with life insurance and annuity sales in a given month. The duration represents the maturity of the contract. For whole life products, it is assumed that the insured would die at 85 years old.

Panel C reports the characteristics of new contracts sold by bank agencies. The markup is the ratio of the present value of expected premium revenue to the present value of expected policy payment, considering all mortality rates. The Residual Markup stands for the residuals of the regression in Table 2.8, discussed in Section 2.2.1. Given that there are some missing values in buyer income, we substitute the missing values with the median income of the same age and gender group in the same city. The Delta/Value stands for the mortality delta following Koijen, Van Nieuwerburgh and Yogo (2016) at the point when the insurance is sold, divided by the insurance value.

Table 2.1: Summary statistics

(a) All SalesmanID (Branch) Obs

year	# of New Contracts	Proportion Life	Proportion Annuity	Proportion A&H	Premium (New) (Million RMB)	Proportion Life	Proportion Annuity	Proportion A&H
2009	102348	52.66%	14.11%	33.24%	1657.35	88.35%	10.97%	0.68%
2010	155802	36.14%	6.67%	57.19%	2126.73	92.50%	5.96%	1.54%
2011	68574	63.69%	13.44%	22.87%	1629.86	93.20%	5.58%	1.23%
2012	65845	51.74%	14.39%	33.87%	1349.33	85.36%	11.53%	3.10%
2013	53832	51.85%	14.56%	33.60%	1037.21	88.60%	8.75%	2.65%
2014	62999	35.00%	15.87%	49.14%	1228.16	80.05%	16.31%	3.64%
2015	61173	22.45%	36.99%	40.56%	1357.89	39.18%	57.99%	2.83%
2016	86899	24.16%	45.29%	30.55%	1541.83	32.93%	64.90%	2.18%

(b) Branch-Month Obs. (Bank Agency)

	N	mean	std	p5	median	p95
# of New Contracts	119515	0.86	1.61	0.00	0.00	3.00
Total Premium	119515	53.76	165.48	0.11	15.00	210.00
Qualified Premium	119515	16.30	81.07	0.00	0.06	65.66
New Premium	55517	83.78	229.13	0.10	340.00	20.00
New Annuity Premium	55517	18.45	109.67	0.00	90.00	0.00
New Life Premium	55517	65.20	201.87	0.00	290.00	10.00
Lapsation	55517	0.01	0.09	0.00	0.00	0.00
Life Duration	38685	7.41	7.34	5.00	6.00	10.00
Annuity Duration	12281	20.65	15.51	10.00	12.50	55.00

(c) New Contracts (Bank Agency)

	N	Mean	Std. Dev.	p5	Median	p95
Markup	146310	1.085	0.119	0.941	1.093	1.221
Residual Markup	146310	0.000	0.039	-0.058	0.001	0.054
Male Ratio	146310	0.395	0.489	0.000	0.000	1.000
Buyer Age	146310	48.722	13.560	26.000	49.000	70.000
Buyer Income (K RMB)	146310	70.829	92.015	30.000	50.000	140.000
Delta/Value	141513	0.180	0.201	0.038	0.101	0.564

Table 2.2: New contract selling due to binding constraints

This table reports the estimated regression coefficients β_1 and β_2 from

$$\begin{aligned} y_{i,t} = \quad & \beta_1 D_{QR^{L2}_{i,t}<C_{i,t}} \times (QR^{L2}_{i,t} - C_{i,t}) + \beta_2 D_{2014} \times D_{QR^{L2}_{i,t}<C_{i,t}} \times (QR^{L2}_{i,t} - C_{i,t}) \\ & + \sum_j \gamma_{1,j} D_j \times (QR^{L2}_{i,t} - C_{i,t}) + \sum_j \gamma_{2,j} D_j D_{QR^{L2}_{i,t}<C_{i,t}} + \theta_t + \eta_i + \epsilon_{i,t}, \end{aligned}$$

where i stands for the bank branch and t stands for the year-quarter. The observations are in the branch-month level. $QR_{i,t}$ is the qualified ratio in the previous two months, which is heterogeneous for each branch i. Institution ID. $C_{i,t}$ is the city-level bank-specific threshold, namely the required qualified ratio in the previous four quarters defined in the data section, which is identical for all branches at each quarter with the same Institution ID. D_{2014} is a dummy for the post–2014 period. We allow the coefficient of the single terms $\{(QR^{L2}_{i,t} - C_{i,t}), D_{QR^{L2}_{i,t}<C_{i,t}}\}$ to have different coefficients $\{\gamma_{1,j}, \gamma_{2,j}\}$ in each year. The time-fixed effect θ_t is captured at the year-quarter level. The panel fixed effect η_i is captured at the bank branch level. All standard errors are double-clustered at year-quarter and branch level. T-stats are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

In Panel A, $y_{i,t}$ is set to be the qualified premium ratio from all new contracts minus the branch's qualified new premium ratio in the last four quarters. In the first column, only quarter-end month observations are included. In the second column, only observations in the first and second month within a quarter are included. In the last column, the differences of coefficients in the first two columns are reported, by running a pooling regression and introducing difference dummies. The coefficients change from $(\beta_1, \beta_2, \{\gamma_{s,j}\}_{1,2}, \theta_t, \eta_i)$ to $(\beta_1 + \delta^*\beta_{\delta,1}, \beta_2 + \delta^*\beta_{\delta,2}, \{\gamma_{s,j} + \delta^*\gamma_{\delta,s,j}\}_{1,2}, \theta_t + \delta^*\theta_{\delta,t}, \eta_i + \delta^*\eta_{\delta,i})$. Here, $\delta^* = 1$ for all sample month observations and 0 otherwise. The reported coefficients in the last column are $(\beta_{\delta,1}, \beta_{\delta,2}, \{\gamma_{\delta,s,j}\}_{1,2}, \theta_{\delta,t}, \eta_{\delta,i})$.

Panel B reports the regression results using the ratio of qualified premiums from new life insurance to the total new premium, subtracting this ratio in the last four quarters of a given branch as $y_{i,t}$. All other settings are identical to Panel A.

Panel C reports the regression results using the ratio of qualified premiums from new annuities to the total new premium, subtracting this ratio in the last four quarters of a given branch as $y_{i,t}$. All other settings are identical to Panel A.

Table 2.2: New contract selling due to binding constraints

(a) Dependent Variable:Qualified Ratio from New contracts - Last Four Qtrs' New Ratio

	Sample Months	Placebo Months	Diff.
β_1	0.082	-0.172	0.254
	(0.48)	(-1.47)	(1.25)
β_2	**-0.528****	**0.206**	**-0.735*****
	(-2.42)	**(1.20)**	**(-2.71)**
Obs.	6117	14740	20857
R^2	0.43	0.337	0.368

(b) Dependent Variable: (New Qualified Life/Total New Premium) - New Qualified Life Ratio Cutoff

	Sample Months	Placebo Months	Diff.
β_1	-0.019	-0.084	0.066
	(-0.12)	(-0.70)	(0.35)
β_2	**0.174**	**0.175**	**-0.001**
	(0.81)	**(1.19)**	**(-0.01)**
Obs.	5807	14425	20232
R^2	0.362	0.314	0.329

(c) Dependent Variable: (New Qualified Annuity/Total New Premium) - New Qualified Annuity Ratio Cutoff

	Sample Months	Placebo Months	Diff.
β_1	0.07	-0.062	0.133
	(0.66)	(-1.07)	(1.17)
β_2	**-0.700*****	**0.011**	**-0.711*****
	(-4.17)	**(0.09)**	**(-3.56)**
Obs.	5807	14425	20232
R^2	0.483	0.343	0.393

Table 2.3: Qualified premium revenue due to binding constraints

This table reports the estimated regression coefficients β_1 and β_2 for the panel regression model identical to Table 2.2. In Panel A, $y_{i,t}$ uses the ratio of the total qualified premium to the total premium for each month of a given bank branch, subtracting this ratio calculated using the premium from the last four quarters. The qualified premium consists of the qualified premium from new contracts and also the qualified premium from old contracts. In Panel B, $y_{i,t}$ uses the ratio of all qualified premiums from life insurance contracts, also subtracting this ratio calculated using the premium from the last four quarters. Panel C substitutes $y_{i,t}$ into the ratio of all qualified premiums from annuities to total premium, also subtracting this ratio in the last four quarters. All the other variable settings are identical to Table 2.2. T-stats are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

(a) Dependent Variable: Qualified Ratio
- Last Four Qtrs' Ratio

	Sample Months	Placebo Months	Diff.
β_1	0.086	-0.061	0.147
	(0.92)	(-0.80)	(1.24)
β_2	**-0.520*** **	**-0.065**	**-0.455*** **
	(-4.28)	**(-0.58)**	**(-2.79)**
Obs.	15075	31923	46998
R^2	0.453	0.317	0.362

(b) Dependent Variable: (Qualified Life/Total Premium)
- Qualified Life Ratio Cutoff

	Sample Months	Placebo Months	Diff.
β_1	0.075	-0.017	0.091
	(0.87)	(-0.21)	(0.79)
β_2	**-0.238***	**-0.09**	**-0.148**
	(-2.01)	**(-0.85)**	**(-0.95)**
Obs.	14688	31629	46317
R^2	0.373	0.265	0.3

(c) Dependent Variable: (Qualified Annuity/Total Premium)
- Qualified Annuity Ratio Cutoff

	Sample Months	Placebo Months	Diff.
β_1	-0.011	-0.058	0.047
	(-0.21)	(-1.19)	(0.67)
β_2	**-0.286*** **	**0.043**	**-0.329*** **
	(-4.80)	**(0.60)**	**(-3.64)**
Obs.	14688	31629	46317
R^2	0.336	0.249	0.279

Table 2.4: Random cutoffs

This table reports the estimated regression coefficients β_1 and β_2 for the panel regression model similar to Table 2.2. The dependent variable $y_{i,t}$ is identical to the Panel A of Table 2.2. The only difference is that $C_{i,t}$ are substituted by $C^*_{i,t} = 0.2 + \sigma\epsilon^{'}_{i,t}$, where $\epsilon^{'}_{i,t}$ is generated from $N(0,1)$ and σ is set to be 0.1 in Panel A, 0.2 in Panel B, or 0.3 in Panel C. All other settings are identical to Table 2.2. T-stats are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

(a) 0.1 Std. White Noise in Cutoff

	Sample Months	Placebo Months	Diff.
β_1	-0.081	-0.177*	0.095
	(-0.72)	(-1.85)	(0.66)
β_2	**-0.315***	**0.143**	**-0.459****
	(-1.86)	**(0.96)**	**(-2.09)**
Obs.	6117	14740	20857
R^2	0.424	0.336	0.364

(b) 0.2 Std. White Noise in Cutoff

	Sample Months	Placebo Months	Diff.
β_1	-0.243**	-0.158***	-0.085
	(-2.42)	(-2.93)	(-0.75)
β_2	**-0.231**	**0.093**	**-0.324***
	(-1.46)	**(0.81)**	**(-1.68)**
Obs.	6117	14740	20857
R^2	0.41	0.327	0.354

(c) 0.3 Std. White Noise in Cutoff

	Sample Months	Placebo Months	Diff
β_1	-0.191***	-0.196***	0.004
	(-3.17)	(-3.67)	(0.06)
β_2	**-0.136**	**0.115**	**-0.252***
	(-1.29)	**(1.20)**	**(-1.83)**
Obs.	6117	14740	20857
R^2	0.401	0.319	0.346

Table 2.5: Cosmetic change in old contracts

This table reports the estimated regression coefficients β_1 and β_2 for the panel regression model identical to Table 2.2. The dependent variable $y_{i,t}$ is the ratio of delayed unqualified premiums to all scheduled premiums from existing contracts in a given month, also substracting this ratio in the previous four quarters. The delayed unqualified premium ratio is

$$\text{Delayed Unqualified Ratio} = \frac{\text{Delayed scheduled unqualified premium}}{\text{Total scheduled unqualified premium}},$$

and the total scheduled unqualified premium is the sum of delayed scheduled unqualified premium and observed unqualified premium. All other settings are identical to Table 2.2. T-stats are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Sample Months	Placebo Months	Diff.
β_1	0.238*	0.224***	0.014
	(1.78)	(3.01)	(0.09)
β_2	**-0.440****	**-0.153**	**-0.287**
	(-2.74)	**(-1.38)**	**(-1.47)**
Obs.	5615	11762	17377
R^2	0.355	0.241	0.278

Table 2.6: Overall premium revenue

This table reports the estimated regression coefficients β_1 and β_2 for the panel regression model identical to Table 2.2. In Panel A, the dependent variable $y_{i,t}$ is the logarithm of the total premium from the new contract selling of the branch in a given month. It is changed to the total premium from the new life insurance (annuity) contracts in Panel B (C). All other settings are identical to Table 2.2. T-stats are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

(a) Dependent Variable: Log(Total New Premium+1)

	Sample Months	Placebo Months	Diff.
β_1	0.266	0.172	0.075
	(0.57)	(1.47)	(0.12)
β_2	**-0.604**	**-0.206**	**-0.651**
	(-0.81)	**(-1.20)**	**(-0.70)**
Obs.	6187	14740	21066
R^2	0.411	0.337	0.411

(b) Dependent Variable: Log(Life New Premium+1)

	Sample Months	Placebo Months	Diff.
β_1	0.526	0.468	0.058
	(0.44)	(0.56)	(0.04)
β_2	**4.777****	**0.065**	**4.712***
	(2.44)	**(0.04)**	**(1.95)**
Obs.	6187	14879	21066
R^2	0.597	0.543	0.56

(c) Dependent Variable: Log(Annuity New Premium+1)

	Sample Months	Placebo Months	Diff.
β_1	0.507	-0.795	1.303
	(0.46)	(-1.30)	(1.06)
β_2	**-7.193*****	**0.307**	**-7.499*****
	(-4.27)	**(0.25)**	**(-3.67)**
Obs.	6187	14879	21066
R^2	0.674	0.623	0.639

Table 2.7: Lapsation rates

This table reports the estimated regression coefficients β_1 and β_2 for the panel regression model identical to Table 2.2. In Panel A (B), the dependent variable $y_{i,t}$ is the equal-weighted (value-weighted) lapsation rate of new contracts sold by the branch in a given month. The lapsed contract is defined as the contract lapses for personal reasons within the next 12 months. All other settings are identical to Table 2.2. T-stats are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

(a) Dependent Variable: Avg. Lapsation Rate (Equal Weighted)

	Sample Months	Placebo Months	Diff.
β_1	0.013	0.006	0.007
	(0.49)	(0.39)	(0.25)
β_2	**-0.021**	**-0.006**	**-0.015**
	(-0.61)	**(-0.29)**	**(-0.38)**
Obs.	6187	14879	21066
R^2	0.301	0.167	0.214

(b) Dependent Variable: Avg. Lapsation Rate (Value Weighted)

	Sample Months	Placebo Months	Diff.
β_1	0.012	-0.009	0.022
	(0.42)	(-0.46)	(0.63)
β_2	**-0.013**	**0.007**	**-0.02**
	(-0.33)	**(0.29)**	**(-0.44)**
Obs.	6187	14879	21066
R^2	0.304	0.163	0.21

Table 2.8: Markup regression

This table reports the estimated regression coefficients for the panel regression model

$$Markup_{j,k,t} = \gamma_1 \mathbf{X}'_{\mathbf{j,t}} + \gamma_2 \mathbf{Z}'_{\mathbf{j,t}} + \gamma_3 \text{vec}(\mathbf{X}'_{\mathbf{j,t}} \mathbf{Z}_{\mathbf{i,t}}) + \theta_t + \eta_i + \epsilon_{i,t},$$

where k is the insurance products insured the person j in the branch i at the quarter t. $Markup_{j,k,t}$ is based on the ratio of the expected claim amount within the contract maturity to the expected present value of the total premium. $X_{j,t}$ represents insured characteristics, including the logarithm of annual income, a male dummy variable, and age. $Z_{j,t}$ is a vector of product characteristics, consisting of the duration and whether the product pays dividends. The time-fixed effect θ_t is captured at the year-quarter level. The panel fixed effect η_i is captured at the bank branch level. In the first three columns, the observations only consist of life insurance products. The first column uses the whole sample, while the second (third) column uses the sub-sample of contracts sold after (before) the regulation reform date. In the last three columns, the observations only consist of annuity products. These three columns follow the identical sub-sample division rule as the first three columns. All standard errors are double-clustered at the year-quarter and branch levels. T-stats are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Life Insurance			**Annuity**		
	Whole	After 2014	Before 2014	Whole	After 2014	Before 2014
Income	0.010*	0.005	0.001	-0.023**	-0.026**	0.014
	(1.92)	(0.89)	(0.03)	(-2.63)	(-2.71)	(1.11)
Dividend	(0.03)	0.039*	-0.133*	0.283***	0.258***	0.00
	(-1.41)	(2.02)	(-1.95)	(4.61)	(3.56)	(0.00)
Gender	-0.01	0.02	0.06	0.059***	0.085***	-0.02
	(-1.12)	(1.62)	(1.05)	(3.65)	(5.24)	(-0.91)
Age	-0.004***	-0.004***	-0.009***	0.00	0.002*	0.00
	(-11.02)	(-6.14)	(-4.69)	(1.66)	(1.78)	(-0.37)
Duration	0.004***	0.006**	0.002***	0.007***	0.007***	0.010***
	(2.73)	(2.52)	(4.95)	(11.47)	(9.93)	(12.07)
Income*Dividend	-0.015***	-0.017***	-0.01	0.025***	0.027**	0.00
	(-3.86)	(-3.65)	(-0.57)	(3.49)	(2.63)	(0.00)
Gender*Dividend	0.024***	0.020***	-0.07	-0.027***	-0.035**	0.00
	(5.20)	(3.13)	(-1.15)	(-2.68)	(-2.67)	(0.00)
Age*Dividend	0.002***	0.001***	0.007***	-0.003***	-0.003**	0.00
	(7.38)	(3.29)	(3.68)	(-2.82)	(-2.27)	(0.00)
Income*Duration	0.00	0.00	0.008***	0.008**	0.010**	0.00
	(0.90)	(0.78)	(3.34)	(2.16)	(2.46)	(-0.15)
Gender*Duration	-0.022***	-0.045***	0.00	-0.027***	-0.037***	0.01
	(-3.73)	(-4.68)	(-0.44)	(-3.40)	(-4.73)	(0.47)
Age*Duration	0.001***	0.00	0.001***	-0.001**	-0.001**	0.00
	(4.55)	(0.87)	(5.48)	(-2.04)	(-2.42)	(0.62)
Obs.	69787	26605	41972	6769	5784	759
Adjusted R^2	0.729	0.753	0.592	0.737	0.695	0.825

Table 2.9: Binding constraints on the markup and markup residual

This table reports the estimated regression coefficients β_1 and β_2 for the panel regression model similar to Table 2.2 but uses contract-level observations. The $y_{i,t}$ in Panel A is the markup of each contract, and the $y_{i,t}$ in Panel B is the regression residual of the markup. These residuals are from regressions identical to Table 2.8 but conducted year-by-year and separately according to annuity or life insurance. In the first three columns of each panel, all contracts are included. In the middle three columns, we focus only on qualified contracts. The last three columns take the remaining unqualified sub-sample contracts. All independent and control variables are identical to Table 2.2 and at the branch level. The regressions use weighted OLS based on the contract's expected premium revenue. All standard errors are still double-clustered at year-quarter and branch level. T-stats are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

(a) Dependent Variable: Markup

	All			Qualified			Unqualified		
	Sample	Placebo	Diff.	Sample	Placebo	Diff.	Sample	Placebo	Diff.
β_1	-0.002	-0.031	0.029	0.037	-0.038	0.075	0.007	0.018	-0.011
	(-0.05)	(-1.37)	(0.69)	(0.66)	(-1.13)	(1.21)	(0.37)	(1.14)	(-0.48)
β_2	**0.007**	**0.022**	**-0.015**	**-0.048**	**0.000**	**-0.048**	**-0.001**	**-0.019**	**0.018**
	(0.12)	**(0.74)**	**(-0.22)**	**(-0.37)**	**(-0.00)**	**(-0.34)**	**(-0.05)**	**(-1.04)**	**(0.64)**
Obs.	11508	29695	41203	1996	4365	6361	9185	24952	34137
R^2	0.88	0.868	0.873	0.883	0.86	0.868	0.825	0.869	0.862

(b) Dependent Variable: Residual Markup

	All			Qualified			Unqualified		
	Sample	Placebo	Diff.	Sample	Placebo	Diff.	Sample	Placebo	Diff.
β_1	-0.01	0.007	-0.017	-0.012	-0.011	0.00	-0.018	0.010	-0.028
	(-0.63)	(0.56)	(-0.87)	(-0.65)	(-0.59)	(-0.01)	(-0.99)	(0.67)	(-1.20)
β_2	**-0.004**	**-0.002**	**-0.002**	**-0.038**	**0.001**	**-0.039**	**-0.002**	**0.001**	**-0.004**
	(-0.14)	**(-0.11)**	**(-0.07)**	**(-0.77)**	**(0.04)**	**(-0.67)**	**(-0.12)**	**(0.08)**	**(-0.14)**
Obs.	11508	29695	41203	1996	4365	6361	9185	24952	34137
R^2	0.32	0.24	0.268	0.427	0.341	0.374	0.317	0.232	0.26

Table 2.10: Binding constraints on investor's characteristics

This table reports the estimated regression coefficients β_1 and β_2 for the panel (Probit) regression model identical to Table 2.9. The $y_{i,t}$ in Panel A is the buyer male dummy, and the regressions are based on Probit model. The $y_{i,t}$ in Panel B is the buyer age, and the $y_{i,t}$ is the logarithm of the buyer's annual income. All standard errors are still double-clustered at year-quarter and branch level. T-stats are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

(a) Dependent Variable: Buyer Gender

	All			Qualified			Unqualified		
	Sample	Placebo	Diff.	Sample	Placebo	Diff.	Sample	Placebo	Diff.
β_1	-0.032	0.289	-0.339	-0.844	-0.324	-0.428	0.400	0.604	-0.286
	(-0.07)	(0.72)	(-0.60)	(-1.32)	(-0.63)	(-0.63)	(0.78)	(1.27)	(-0.41)
β_2	**0.359**	**-0.446**	**0.806**	**0.672**	**0.180**	**0.510**	**-0.241**	**-0.850***	**0.617**
	(0.69)	**(-0.93)**	**(1.15)**	**(0.95)**	**(0.20)**	**(0.45)**	**(-0.33)**	**(-1.77)**	**(0.72)**
Obs.	12428	30429	42857	2490	4832	7322	9938	25597	35535
Pseudo R^2	0.0131	0.0022	0.0055	0.0578	0.004	0.0241	0.0075	0.0031	0.0044

(b) Dependent Variable: Buyer Age

	All			Qualified			Unqualified		
	Sample	Placebo	Diff.	Sample	Placebo	Diff.	Sample	Placebo	Diff.
β_1	-0.117	2.696	-2.813	-2.686	-3.428	0.742	3.454	4.232	-0.778
	(-0.04)	(0.75)	(-0.58)	(-1.08)	(-0.64)	(0.13)	(0.87)	(0.87)	(-0.12)
β_2	**1.408**	**-0.468**	**1.876**	**1.462**	**10.881**	**-9.419**	**0.374**	**-5.138**	**5.512**
	(0.34)	**(-0.12)**	**(0.33)**	**(0.26)**	**(1.46)**	**(-1.03)**	**(0.05)**	**(-0.96)**	**(0.62)**
Obs.	11508	29695	41203	1996	4365	6361	9185	24952	34137
R^2	0.445	0.351	0.378	0.59	0.582	0.585	0.414	0.311	0.339

(c) Dependent Variable: Log Buyer Income

	All			Qualified			Unqualified		
	Sample	Placebo	Diff.	Sample	Placebo	Diff.	Sample	Placebo	Diff.
β_1	0.617	0.520	0.097	1.082**	0.612	0.470	0.189	-0.590	0.779
	(1.39)	(1.62)	(0.18)	(2.54)	(1.68)	(0.86)	(0.27)	(-1.28)	(0.99)
β_2	**0.188**	**-0.240**	**0.428**	**-0.458**	**-0.453**	**-0.005**	**0.955**	**0.791**	**0.164**
	(0.33)	**(-0.68)**	**(0.66)**	**(-0.68)**	**(-0.91)**	**(-0.01)**	**(1.15)**	**(1.52)**	**(0.18)**
Obs.	4502	11577	16079	1410	3248	4658	2829	8052	10881
R^2	0.709	0.668	0.68	0.608	0.666	0.65	0.836	0.705	0.739

Table 2.11: New contracts for PA channel

This table reports the estimated regression coefficients β_1 and β_2 for the panel regression models similar to Table 2.2 but use the personal agency observations. The observations are at the SalesmanID level, corresponding to the bank branches. The virtual binding constraints are set at the InstitutionID level, corresponding to the bank of each city. The three panels are identical to the three panels in Table 2.2 respectively. All standard errors are still double-clustered at year-quarter and branch level. T-stats are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

(a) Dependent Variable:Qualified Ratio from New contracts - Last Four Qtrs' New Ratio

	Sample Months	Placebo Months	Diff.
β_1	0.018	-0.035	0.052
	(0.12)	(-0.36)	(0.31)
β_2	**0.028**	**0.078**	**-0.05**
	(0.18)	**(0.79)**	**(-0.30)**
Obs.	4106	8146	12252
R^2	0.757	0.586	0.663

(b) Dependent Variable: (New Qualified Life/Total New Premium) - New Qualified Life Ratio Cutoff

	Sample Months	Placebo Months	Diff.
β_1	0.246	-0.071	0.317
	(0.89)	(-0.62)	(1.10)
β_2	**0.04**	**0.205***	**-0.165**
	(0.14)	**(1.73)**	**(-0.55)**
Obs.	4106	8146	12252
R^2	0.534	0.558	0.551

(c) Dependent Variable: (New Qualified Annuity/Total New Premium) - New Qualified Annuity Ratio Cutoff

	Sample Months	Placebo Months	Diff.
β_1	-0.228	0.036	-0.265
	(-0.66)	(0.22)	(-0.72)
β_2	**-0.012**	**-0.127**	**0.115**
	(-0.03)	**(-0.76)**	**(0.30)**
Obs.	4106	8146	12252
R^2	0.565	0.555	0.558

Table 2.12: Other tests for PA channel

This table reports the estimated regression coefficients β_1 and β_2 for the panel regression models indetical to Table 2.11 using the personal agency observations. Panel A and Panel B are identical to Panel A of Table 2.6 and Panel A of Table 2.7 respectively. All standard errors are still double-clustered at year-quarter and branch level. T-stats are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

(a) Dependent Variable: Log(Total New Premium)

	Sample Months	Placebo Months	Diff.
β_1	0.102	0.203	0.203
	(0.12)	(0.40)	(0.40)
β_2	**-0.134**	**-0.271**	**-0.271**
	(-0.14)	**(-0.50)**	**(-0.51)**
Obs.	4137	8215	12352
R2	0.696	0.671	0.68

(b) Dependent Variable: Lapsation Rate (Value Weighted)

	Sample Months	Placebo Months	Diff.
β_1	0.068*	-0.041	0.109**
	(1.92)	(-1.63)	(2.58)
β_2	**-0.06**	**0.027**	**-0.086***
	(-1.55)	**(0.99)**	**(-1.89)**
Obs.	4137	8215	12352
R2	0.184	0.217	0.206

Bibliography

Egan, Mark, Gregor Matvos, and Amit Seru. 2019. "The market for financial adviser misconduct." *Journal of Political Economy*, 127(1): 233–295.

Ge, Shan. 2022. "How do financial constraints affect product pricing? evidence from weather and life insurance premiums." *The Journal of Finance*, 77(1): 449–503.

Ge, Shan, and Michael S Weisbach. 2021. "The role of financial conditions in portfolio choices: The case of insurers." *Journal of Financial Economics*, 142(2): 803–830.

Koijen, Ralph SJ, and Motohiro Yogo. 2015. "The cost of financial frictions for life insurers." *American Economic Review*, 105(1): 445–475.

Koijen, Ralph SJ, Stijn Van Nieuwerburgh, and Motohiro Yogo. 2016. "Health and mortality delta: Assessing the welfare cost of household insurance choice." *The Journal of Finance*, 71(2): 957–1010.

2.A Appendix

Table 2.A.1: New premium responses to qualified ratios across different cities

This table reports the estimated regression coefficients β_1 and β_2 for the panel regression models identical to Panel A of Table 2.2. Each panel use the subsample observations based on the cities. Panel A consists of tier 1 cities including Beijing, Shanghai and Guangzhou. Panel B includes observations in Nanjing, Chengdu and Wuhan as the tier 2 cities. Panel C focuses on observations in Zhengzhou, Lanzhou and Shenyang, which are tier 3 cities. All standard errors are still double-clustered at year-quarter and branch level. T-stats are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

(a) Tier 1 Cities

	Sample Months	Placebo Months	Diff.
β_1	-0.01	-0.207	0.198
	(-0.04)	(−1.43)	(0.73)
β_2	**-0.492***	**0.212**	**-0.703****
	(-1.76)	**(1.01**	**(-2.08)**
Obs.	4078	10185	14263
R^2	0.433	0.365	0.387

(b) Tier 2 Cities

	Sample Months	Placebo Months	Diff.
β_1	0.425	-0.183	0.608*
	(1.55)	(-0.90)	(1.82)
β_2	**-0.826***	**0.195**	**-1.021****
	(-1.98)	**(0.78)**	**(-2.17)**
Obs.	1333	2846	4179
R^2	0.413	0.336	0.362

(c) Tier 3 Cities

	Sample Months	Placebo Months	Diff.
β_1	-0.267	0.09	-0.356
	(-0.99)	(0.30)	(-0.92)
β_2	**0.632**	**-0.078**	**0.71**
	(1.33)	**(-0.21)**	**(1.28)**
Obs.	706	1709	2415
R^2	0.601	0.414	0.477

Table 2.A.2: Total premium responses to qualified ratios across different cities

This table reports the estimated regression coefficients β_1 and β_2 for the panel regression models identical to Panel A of Table 2.3. Each panel use the subsample observations based on the cities. Panel A consists of tier 1 cities including Beijing, Shanghai and Guangzhou. Panel B includes observations in Nanjing, Chengdu and Wuhan as the tier 2 cities. Panel C focuses on observations in Zhengzhou, Lanzhou and Shenyang, which are tier 3 cities. All standard errors are still double-clustered at year-quarter and branch level. T-stats are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

(a) Tier 1 Cities

	Sample Months	Placebo Months	Diff.
β_1	0.064	-0.12	0.184
	(0.69)	(-1.39)	(1.47)
β_2	**-0.511*** **	**0.03**	**-0.542*** **
	(-4.27)	**(0.23)**	**(-3.05)**
Obs.	9904	21205	31109
R^2	0.456	0.338	0.377

(b) Tier 2 Cities

	Sample Months	Placebo Months	Diff.
β_1	0.489***	0.217*	0.272
	(3.02)	(1.73)	(1.32)
β_2	**-0.968*** **	**-0.410** **	**-0.558** **
	(-5.07)	**(-2.49)**	**(-2.27)**
Obs.	3174	6504	9678
R^2	0.436	0.292	0.336

(c) Tier 3 Cities

	Sample Months	Placebo Months	Diff.
β_1	-0.252***	-0.196	-0.056
	(-2.92)	(-1.06)	(-0.27)
β_2	**0.058**	**-0.028**	**0.086**
	(0.24)	**(-0.13)**	**(0.27)**
Obs.	1997	4214	6211
R^2	0.515	0.353	0.409

Table 2.A.3: Mortality delta per contract due to the binding constraints

This table reports the estimated regression coefficients β_1 and β_2 for the panel regression models identical to Table 2.2 but with different $y_{i,t}$. The $y_{i,t}$ in this table is the total premium weighted mortality delta, normalized to the insurance value. The mortality delta follows Koijen, Van Nieuwerburgh and Yogo (2016). All standard errors are still double-clustered at year-quarter and branch level. T-stats are reported in parentheses. ***, **, and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Sample Months	Placebo Months	Diff.
β_1	-0.193	0.03	-0.223
	(-1.47)	(0.17)	(-1.08)
β_2	**0.778****	**-0.223**	**1.002****
	(2.23)	**(-0.63)**	**(2.06)**
Obs.	4767	11870	16637
R^2	0.541	0.449	0.477

Chapter 3

Mutual Fund Sector Concentration and Size Effects

Jiaxing Tian[1]

Mutual funds are known to experience diseconomies of scale. This paper examines the relationship between the magnitude of this negative size effect and fund sector concentration. It finds a strong correlation indicating that funds in more concentrated sectors exhibit more severe diminishing returns to scale compared to those in less concentrated sectors. Notably, this relationship follows a monotonic pattern. The paper proposes a potential explanation: in highly concentrated sectors, fund flows are less sensitive to past returns. However, in such sectors, marketing expenses appear to positively influence flow sensitivity to good performance, while showing a neutral effect in response to poor performance. The findings suggest that large funds in concentrated sectors may invest more in marketing efforts, but this does not necessarily translate to better future performance.

[1]This is the revised version of my second-year paper. I thank Dong Lou, Christopher Polk, Michela Verardo, Cameron Peng, Shiyang Huang, Jiahong Shi and seminar participants at LSE for helpful comments. All remaining errors are my own.

3.1 Introduction

Mutual fund size has been proven to be negatively correlated with its performance (Berk and Green, 2004). This could be linked to liquidity constraints (Chen et al., 2004), industry size (Pástor and Stambaugh, 2012), or a mismatch between fund scale and manager's skill (Song, 2017). The mismatch, where money flows into funds with poor performance, has been explained by investors ignoring common factors or by suggesting that size and fund alpha are indeed an entity.[2] However, questions such as how the level of diseconomies of scale varies across fund categories and which kinds of funds are more likely to suffer from negative size effects remain unanswered.

In the realm of mutual funds, investor behavior dynamics play a crucial role in fund performance and size. Investors often exhibit preferences that may not align with optimal investment strategies, leading to peculiar trends such as larger inflows into underperforming funds. This phenomenon, known as diseconomies of scale, poses intriguing questions regarding the factors that drive investors to allocate their capital into larger funds, irrespective of performance. Understanding which categories of funds suffer from diseconomies of scale and which do not is crucial in comprehending the drivers of investor investment strategies toward larger funds. In this paper, I delve into the relationship between mutual fund sector concentration and the manifestation of diseconomies of scale, shedding light on the mechanisms underlying investor decisions in fund selection. In summary, fund competition across different fund sectors plays an important role in investment preferences and thus explains the cross-sectional differences in the diseconomies of scale.

As investors gravitate towards funds with objectives that match their own investment goals, understanding the competitive landscape within funds of similar objectives becomes paramount. For instance, investors focused on large capital stocks may prioritize funds catering to this segment, showing limited concern for the performance of small-cap funds. Therefore, comprehending how funds with aligned objectives vie against each other in terms of characteristics and perfor-

[2] Large mutual funds are more likely to suffer from trading impacts or may ignore managerial skill to achieve relatively safe performance.

mance is essential in deciphering inflow patterns and fund size dynamics.

This paper reveals a noteworthy trend: as sector concentration within the mutual fund landscape intensifies, the magnitude of diseconomies of scale amplifies proportionately. From the very beginning, the market concentration, indicated as the Herfindahl–Hirschman index (HHI), is negatively correlated with sector returns. Those sectors with higher market concentration tend to have lower returns. This effect is more significant in large funds within such concentrated sectors. With a 10% higher concentration, a 10% higher fund size would lead to a 2.1 basis point per month lower return compared to the same large fund in low concentration sectors. This result is robust across all HHI measures.

Such a relationship between sector concentration and the level of diseconomies of scale is monotonic. I run the Fama-Macbeth regressions sector by sector to reveal this monotonic relationship. Funds in the lowest concentrated sectors even exhibit no diseconomies of scale, indicating that large funds and small funds have no significant differences in terms of returns. However, funds in the most concentrated sectors exhibit the highest level of diseconomies of scale.

In highly concentrated sectors, the size effect on fund inflows surges, reaching levels nearly six times higher compared to sectors with lower concentration levels. This monotonic relationship underscores the significant impact of sector concentration on fund size dynamics, suggesting a dominance of combined effects over individual factors, particularly the size effect. Conversely, in sectors characterized by lower concentration levels, funds tend to exhibit comparable sizes, fostering sensitivity to returns as proxies for managerial skill. Empirical findings indicate that funds in concentrated sectors display reduced sensitivity to past returns, both in the short and long term.

This diminished sensitivity underscores the influence of sector concentration on investor behavior, where marketing and non-performance-related factors may override considerations of managerial skill in fund selection. I show that large funds in whichever sector are quite identical in terms of different kinds of characteristics, except the marketing expense, proxied by the actual 12(b)-1. Large funds in highly concentrated sectors are more likely to spend more on marketing, leading investors to be more sensitive to positive performance but not sensitive to poor performance in terms of their flows. One possible explanation is that when sectors are highly concentrated, investors lack an anchor to find comparisons, which makes the large funds spend more efforts in marketing and grow even larger. Due to limited efforts of the management team, their performance is usually poor.

Our investigation also explores alternative hypotheses regarding the relationship between sector concentration and diseconomies of scale. Contrary to the notion of endogenous fund behaviors driving market concentration, our analysis reveals no clear correlation between the largest fund size within each sector and sector concentration level. Moreover, the absence of discernible patterns in fund characteristics across sectors further dispels alternative explanations, suggesting a complex interplay of factors shaping market concentration and fund size.

In summary, our study offers empirical insights into the intricate interplay between sector concentration, fund size dynamics, and investor behavior in the mutual fund landscape. While causality remains elusive, our findings underscore the correlation between sector concentration and the magnitude of diseconomies of scale, highlighting the nuanced factors shaping fund inflows and size dynamics.

My analysis is based on Chen et al. (2004) empirical strategy and variable settings. The most related papers are Pástor, Stambaugh and Taylor (2015), Pástor and Stambaugh (2012), Song (2017), and Huang et al. (2023), which argue that diseconomies of scale in mutual funds are driven by mismatch and industry competition. Pástor and Stambaugh (2012) and Pástor, Stambaugh and Taylor (2015) suggest that as the industry expands, funds' ability to outperform diminishes. This raises questions about whether skill improvement can keep pace with industry-level diseconomies of skill. However, skill is a relative measure, and there may always be some smart managers who outperform, given a finite stock pool. The central question is whether capital flows to genuinely skilled managers or to those who are merely lucky but possess strong marketing capabilities. This paper focuses on the mismatch aspect of why skill fails to yield significant returns and how industry-level diseconomies of scale impact individual funds. Song (2017) argues that investors lack the sense to adjust returns for systematic risks, resulting in capital flows that are not optimal. Funds that are insensitive to systematic risks stand to benefit from this behavior. My paper delves deeper into how much mismatch varies across funds, positing that in addition to the absence of risk factor adjustments, investors may be swayed by misleading fund marketing efforts.

This paper is centered on the area of research in fund size and performance. Yan (2008) also provides evidence to explain diminishing returns to scale, which is linked with liquidity constraints. In contrast, Elton, Gruber and Blake (2012) report there is no such relationship when the reduction in expense ratio for large funds is taken into account. This paper suggests that in more concentrated sectors, since marketing benefits inflows, such a relationship persistently exists and becomes even more severe in magnitude. Large funds would consider the role of marketing in these sectors and not just reduce their expenses. Zhu (2018) argues that size and alpha are indeed an entity. Such an idea also suggests that large size would lead to a mismatch with performance and flow. Bhojraj, Jun Cho and Yehuda (2012) suggests that large family funds have an information advantage to form a positive relationship between size and performance. I have controlled for the family fund size in most quantitative analyses of this paper. Pollet and Wilson (2008) argue that when large funds face diminishing returns to scale, they would diversify their portfolio. Diversification would lead to good performance in the future.

Another area related to this paper is the concentration and competition in mutual funds and holding stocks. Kacperczyk, Sialm and Zheng (2005) focus on the industry concentration of stocks held by mutual funds and its relationship with fund performance. In their paper, managerial skill is evident for managers who hold portfolios concentrated in a few industries. However, it may be another case for concentrated fund sectors as suggested in this paper. Funds in concentrated sectors are more likely to be linked to a mismatch between flow and performance. Industry size could be considered as a proxy for sector competition (Pástor and Stambaugh, 2012). However, industry size may not be enough to explain the magnitude of negative size effects. Hoberg, Kumar and Prabhala (2018) study mutual fund peer competitions in terms of stock characteristics held by funds. They suggest funds facing less competition would generate larger future alpha. To eliminate the similarity and competition effects within each sector, I have shown that the competition levels are not different across sectors. The funds that outperform or get large inflows do not necessarily have a skilled manager.

The rest of this paper is organized as follows. Section 3.2 introduces the data, the sample range, and the variables used in this paper. Section 3.3 provides evidence on how mutual fund performance relates to fund size, fund sector concentration, and the combination effect. Section 3.4 shows that the sensitivity of mutual fund flows to past performance and marketing expense is related to sector concentration. Section 3.5 highlights all findings and concludes.

3.2 Data and Variable Settings

In this section, I describe the mutual fund dataset and each variable I use in the analysis. I rely on actively managed open-ended U.S. equity mutual funds from 1999 to 2019 from the CRSP Survivor Bias-Free U.S. Mutual Fund database. To identify diversified equity funds, first, I select the funds with CRSP Style Codes in EDCL, EDCM, EDCS, EDCI, EDYG, EDYB, EDYH, EDYS, or EDYI. These are also the identifications I use to define fund sectors[3]. Some researchers use Lipper Objective codes as a benchmark to identify equity funds like Zheng, Kacperczyk and Sialm (2007) and Hoberg, Kumar and Prabhala (2018). To be consistent with their results, all funds with Lipper Classification codes **not** in EIEI, LCCE, LCGE, LCVE, MCCE, MCGE, MCVE, MLCE, MLGE, MLVE, SCCE, SCGE, or SCVE have been eliminated from my sample.

The primary distinction between Lipper Classification codes and CRSP Objective codes lies in their treatment of "style" funds. Lipper codes interact each investment style with the level of capital that funds target. For instance, LCGE in Lipper code denotes "Large Capital Growth Funds", while CRSP code utilizes classifications such as EDYG for "Growth Funds" and EDCL for "Large Cap. Funds". The preference for CRSP Objective codes as sector identifiers is based on two considerations. Firstly, the interaction of capital and investment style in Lipper codes may lead to imbalanced sectors due to the high correlation between stock size and book-to-market ratio. Secondly, Lipper codes are only available from 1998

[3] All results are robust if Lipper Objective codes are used to define fund sectors.

onwards. Although the analysis begins in 1999, CRSP Objective codes enable seamless integration with historical sector data, as they are defined over the entire sample period in the CRSP Survivor Bias-Free U.S. Mutual Fund database.

For the purpose of this analysis, I eliminate index funds and exchange-traded funds. All funds with subclasses are aggregated into a single fund using total net assets in each subclass. I also exclude funds with less than one year of age and less than $5 million in total assets. There are 5266 funds remaining after these filters.

The fund characteristics, including total net assets, expense ratio, age, turnover ratio, total loads, and 12b(1) fee, are all obtained from the CRSP mutual fund dataset. The stock holding information is from the CDA/Spectrum database provided by Thompson Financial. To merge these two databases, I rely on the Mutual Fund Links dataset provided by Russ Wermers on Wharton Research Data Services (WRDS). To aggregate subclasses, I also rely on the identification Wharton Financial Institution Center Number (WFICN). Except for TNA, which is the sum of each subclass, all other variables are TNA-weighted averages. The fund family TNA is defined as the aggregate TNA of the same management company, i.e., the Management Company Code in the CRSP mutual fund database. The fund flows from month t to $t+\tau$ of fund i are defined as:

$$Flow_{i,t,t+\tau} = \frac{TNA_{i,t+\tau} - TNA_{i,t} \times (1 + ret_{i,t,t+\tau})}{TNA_{i,t}}, \tag{3.1}$$

where $ret_{i,t,t+\tau}$ is the net return from t to $t+\tau$.

To obtain factor-adjusted returns, I follow the work by Chen et al. (2004). All funds are sorted by their total net assets (TNA) at the end of each month, and five portfolios are formed. The factor loadings are estimated using the entire sample of each portfolio. I employ three-factor models in gross and net returns: CAPM (single market factor model), Fama-French three-factor model (FF3, market, size, and book-to-market ratio), and Fama-French three-factor model together with the momentum factor (FF4). The gross return in each month is calculated using the net return in the CRSP database plus one-twelfth of the expense ratio of that fiscal year.

In this paper, I use sector concentration to explain why the magnitude of diseconomies of scale varies from fund to fund. The sector concentration is defined as the Herfindahl-Hirschman Index (HHI) in each CRSP Objective Code sector K:

$$HHI_{K,t} = \sum_{s \in K} \frac{TNA_{s,t}}{\sum_{j \in K} TNA_{j,t}}, \tag{3.2}$$

where $TNA_{s,t}$ is the total net assets of fund s which belongs to sector K in month t. The sector with high HHI has a high concentration, i.e. there are some relatively large funds in the sector.

3.2.1 Descriptive Statistics

Table 3.1 presents summary statistics for the dataset. The analysis is based on fund-month observations, with 1,907 funds in the first month of 1999 and 2,740 funds in the last month of 2019. All fund-month observations have an average total net assets (TNA) of 1.35 billion dollars and family TNA of 78.76 billion dollars. The largest fund owns more than 24 billion in TNA, while the smallest has just above 6 million. The funds have an average 12-month flow of 21.5% and an average age of 10 years. The turnover ratio is 68%, while the total load is around 1.4%. Given the significant variation in fund characteristics, all variables are winsorized from the 1% to 99% level across the entire history in the following analysis. The Herfindahl-Hirschman Index (HHI) ranges from 1.35% (EDYG) to 79.16% (EDYS).

3.2.2 Summary Statistics by Sector

Table 3.2 summarizes additional characteristics for each fund sector. One concern regarding whether market concentration impacts diseconomies of scale is whether large funds in the entire sample are in highly concentrated sectors. According to this alternative hypothesis, it is the fund size itself that matters rather than concentration, implying that the performance of large funds worsens monotonically. Another concern is whether the concentration is an endogenous outcome of stock-holding characteristics affecting returns. Since funds within the same sector share similar objectives, they are expected to hold similar stocks. Any significant deviation in equity holdings among funds within a sector could lead to size dispersion and high concentration. However, Table 3.2 provides evidence against such hypotheses.

In all nine sectors, EDYS has the smallest number of funds, approximately 10, resulting in the largest HHI. Among sectors with more than 50 funds, EDCL, EDYH, and EDYI exhibit large HHIs, while EDYG and EDCS have smaller ones. However, it is not necessarily the case that more funds lead to smaller HHIs. For instance, EDYB has 491 funds but a HHI of 4.58%, whereas EDCM has fewer funds (287) and a smaller HHI (2.65%). Furthermore, large HHIs do not always imply the presence of large funds. In the EDCL sector, for instance, the HHI is 17.30%, but the largest fund has a TNA of only 159 billion. Conversely, in the EDYB sector with a HHI of 4.58%, the largest TNA is 227 billion. These findings suggest that the largest funds are not concentrated in specific sectors, although some sectors have fewer funds, and there are large funds in those sectors as well. The proportions of TNA for the largest, top 3, top 5, and top 10 funds demonstrate that when sectors have similar sizes, HHI is determined by the magnitude of the largest funds within the sector.

Another consideration is whether fund sizes are distributed uniformly across each sector. If the distributions differ among sectors, it could be the case that some funds, either too small or too concentrated in the middle range, perform relatively better. However, this is not observed in my analysis, as the mean absolute deviation (MAD), scaled mean absolute deviation (MAD/Mean), and scaled standard deviation (SD TNA/MEAN) exhibit similar patterns across sectors.

The tracking error and spatial distance provide evidence against the hypothesis that concentration serves as a proxy for sector fixed effects. To compute the tracking error of each fund, I utilize index returns from the CRSP database. The tracking error is determined by the standard deviations of $(ret_{i,t} - ret_{\text{index},t})$ over the past 12 months. The reported values represent historical averages for each month. Table 3.3 presents the matches between sectors and their corresponding tracking indexes. However, a notable issue arises with the EDYS sector, where I employ the US Total Market Index as the benchmark for funds with dedicated short targets, resulting in a significant tracking error for this sector. Conversely, for other sectors, tracking errors range from 3% to 5%, with no discernible pattern regarding how market concentration correlates with tracking error.

The spatial distance metric, as defined by Hoberg, Kumar and Prabhala (2018), determines the peer competition of each fund. It is computed as the average within each sector. While the values may differ in magnitude from those in Hoberg, Kumar and Prabhala (2018) due to the normalization of all three dimensions in their study, I adhere to their model $S^{Orth}_{F,3d} = \{zSize, zrB/M_f, zrMOM\}$, where the orthogonal z-scores of size, book-to-market ratio, and momentum of stocks held by funds are utilized to construct the spatial space. Across all sectors, except for the large capital targeting sector (EDCL), the spatial distances remain relatively consistent, ranging from 0.7 to 1.06. In the large capital sector, it is reasonable to expect fewer large stocks, resulting in fund holdings with similar characteristics. Furthermore, all subsequent results have been demonstrated to be robust even with the elimination of small sectors (EDYS, EDCI) or the large capital sector (EDCL).

3.3 Market Concentration on Diseconomies of Scale

In this section, I initially investigate whether diseconomies of scale hold in my sample and sample period, and whether market concentration contributes to cross-sectional return differences. By interacting size and market concentration, I then determine that the combined effect of size and sector concentration dominates the size effect. Additionally, I observe that large size does not necessarily equate to poor performance in low-concentration sectors, but this effect magnifies significantly in the most highly concentrated sectors.

Following Chen et al. (2004), Table 3.4 and 3.5 present the Fama-Macbeth regression results from 1999-2019 for all funds and sectors in my sample. In addition to fund size and relevant controls, market concentration is included in the regression to test whether concentration itself contributes to cross-sectional returns. Two proxies for market concentration are utilized: the logarithm of $(HHI \times 10000)$, which has the economic interpretation of one basis point change in HHI, and the quintile rank of HHI. Each month, the nine sectors are sorted into five groups based on their HHI. Sectors EDYS and EDYH are excluded until 2010, as the first funds in these sectors started in January 2006 and there are more than 5 funds after 2010. EDYS is always placed in the highest rank group. Subsequently, each sector comprises one or two sectors. EDCI and EDCL consistently fall into the highest rank quintile group, while the ranks of other sectors may vary over time. The groups are rebalanced at the end of each year.

In Table 3.4, the logarithm of HHI is utilized as the proxy for market concentration and added to the regressions. Consistent with previous studies, both gross returns and net returns are tested as dependent variables to demonstrate that expense ratios in the respective years have no impact on size effects. Both types of returns are adjusted by risk factors using four different models. Each column in the table represents a unique adjustment method. In Panel A, the "Raw" column employs raw returns as dependent variables, while in Panel B, "Raw" represents gross returns. In both panels, "Ret-Mkt_Ret" uses raw returns minus market portfolio returns at $t+1$. The subsequent three columns in each panel utilize factor models to adjust returns. "Ret_Beta" applies the CAPM model, while "Ret_FF3" and "Ret_FF4" use the Fama French three-factor model and the three-factor model with the momentum factor, respectively. Following Chen et al. (2004), all factor models are estimated across five fund size groups. At the end of each month, five portfolios are formed based on the TNA of each fund for that month. The entire sample period is utilized to estimate the factor loadings, assuming that each fund within the same size portfolio shares identical loadings.

The regression model is provided below:

$$Ret_{i,t+1} = \mu_t + b_{1,t}\mathrm{LogTNA}_{i,t} + b2_{2,t}\mathrm{HHI}_{i,t} + \gamma_t \boldsymbol{X_{i,t}} + \upsilon_{i,t} \tag{3.3}$$

Eq. 3.3 represents the first stage regression of the Fama-Macbeth procedure in Table 3.4 and Table 3.5, where t denotes the month and i denotes the fund. $\mathrm{LogTNA}_{i,t}$ denotes the logarithm of the TNA of fund i in month t. $\mathrm{HHI}_{i,t}$ represents $Log(HHI \times 10000)$ in Table 3.4 and a dummy variable indicating whether fund i is in the quintile group with the highest concentration in Table 3.5. The control variables include the logarithm of the fund family TNA at month t, the expense ratio of fund i in the last fiscal year up to month t, the previous 12-month flow $Flow_{i,t-12,t-1}$ as defined in Eq. 3.1, the age of the fund in years at the end of month t, the turnover ratio, and the total loads of the last fiscal year up to month t. In the second stage, all t-statistics are adjusted using the fourth-order Newey-West method.

The coefficients of size in Panel A and Panel B exhibit consistency across different columns. The coefficients for gross returns are smaller than those for net returns, with a gap of approximately 0.01, consistent with the findings of Chen et al. (2004). The method of adjusting returns does not affect our results because the size effect pertains to adjusted performance. The reasons why larger funds perform worse are irrelevant to systematic risk. All coefficients on size are statistically significant, with t-statistics exceeding 2.5 in absolute value.

Across these four panels in the two tables, high concentration sectors indicate lower returns in all cases. A one percentage increase in HHI results in a 0.1 basis points decline in future month returns, as shown in Table 3.4. Funds in the highest concentration sectors are expected to have 12.4 basis points lower returns than those in lower concentration sectors, according to Table 3.5. The concentration effect is statistically significant, as indicated by all t-statistics of these coefficients being below -2.5. These results align with the industry size analysis conducted by Pástor, Stambaugh and Taylor (2015).

3.3.1 The Interaction Effect of Size and Market Concentration

One of the key hypotheses posits that market concentration matters for diseconomies of scale. In highly concentrated sectors, diseconomies of scale are expected to become more severe due to the significant mismatch between fund skill and inflows. Conversely, in highly distributed sectors, large funds may be those that genuinely perform well and possess good managerial skills. To demonstrate this, I conduct (first-stage) regressions as shown in Table 3.6.

$$Ret_{i,t+1} = \mu_t + b_{1,t}\text{LogTNA}_{i,t} + b_{2,t}\text{HHI}_{i,t} + \theta_t(\text{LogTNA}_{i,t} \times \text{HHI}_{i,t}) + \boldsymbol{\gamma_t X_{i,t}} + \upsilon_{i,t}. \tag{3.4}$$

To mitigate multicollinearity and its potential bias in coefficient estimations, I incorporate the quintile rank of HHI as the control variable for HHI in Table 3.6. In Panel A, the interaction term utilizes $LogHHI \times LogTNA$, representing the multiplier of size and concentration as shown in Table 3.4. Additionally, I present results using Quintile Rank of HHI $\times$ $LogTNA$ as the interaction term, consistent with the HHI control variable here. All reported returns are net returns.[4]. Control variables remain identical to previous settings in Table 3.4.

In Panel A, the coefficients of size effects become insignificant, as does the sector concentration term. Moreover, the point estimates of these two variables even flip signs. However, the interaction terms are negatively significant, ranging around -0.021, with t-statistics around -2. Consistently, Panel B also exhibits such significant interaction terms and insignificant size and concentration terms. Since Table 3.2 demonstrates that large funds are not necessarily located in highly concentrated sectors, one possible explanation comes to mind.

Market concentration may directly influence the size effect, and together they function as an entity. High concentration sectors are associated with a significant mismatch between skills and inflows. Specifically, higher market concentration may imply fewer investment opportunities or an insignificant "smart money" effect. Second, market concentration may be an endogenous result of large-sized funds in that sector. Funds with marketing power tend to be large and contribute to a concentrated sector. Simultaneously, such marketing power may induce other benefits beyond returns, attracting large inflows. In the following section, I will provide evidence for the first assumption.

Another concern for this analysis is whether the market concentration affects the magnitude of diseconomies of scale monotonically, i.e., whether only small but concentrated sectors like EDYS have such an effect. The monotonic relationship is more likely to support the possible explanation above and rule out the outliers.

[4]The results are similar if gross returns are used, as demonstrated in Table 3.4 and Table 3.5

Table 3.7 shows the results of Fama-Macbeth regression sectors by sectors according to their concentration level using similar settings to Eq. 3.3 without the market concentration term. The sectors are sorted into five groups each month as discussed in Section 3.3.

There is a clear monotonic pattern in Panel A for size effect, except for the second-highest concentrated sectors. In the lowest concentrated sector, the size effect is not significant with a point estimate of -0.023, while it decreases to -0.037 for the larger group and then to -0.052 for the third group. It reaches -0.113 in the largest group. The outlier is the second-highest concentrated group with an insignificant coefficient of -0.024. Sectors EDYH (hedging style) and EDYB (growth and income style) usually lie in this group each month.

One plausible explanation is that although these sectors are highly concentrated, the mismatch in such sectors is not severe. These two investment styles require higher fixed costs to select stocks and also incur high costs to acquire accurate information about hedging or targeting growth together with income. Therefore, positive and long-lasting inflows would more likely mean there is less mismatch because this kind of mismatch would show up directly for those funds with low skills and their returns would fluctuate more. These results are consistent with Panel B when gross returns are used as the dependent variable to eliminate the impact of fees. The second-largest group is still an outlier in my results, although these groups have relatively higher fees (EDYH) or higher loads (EDYB), which is consistent with the explanation that these funds have higher costs. The explanations here are further supported in Section 3.4, where it is shown that the large funds in these two sectors do not exhibit higher marketing expenses. Although concentrated, these sectors require more skills to attract higher inflows from the "smart money".

In conclusion, the magnitude of negative size effects varies across sectors and is consistently related to sector concentration in a monotonic manner. These findings remain robust even when HHIs are calculated using fund family shares instead of individual fund shares. Two hypotheses are proposed to explain this phenomenon: concentration and size effects are an endogenous entity, related to fund characteristics; or concentration exogenously determines the level of size effects. While this paper does not definitively determine which hypothesis is correct, it aims to demonstrate that such comovement may operate through a flow-based channel and could be linked to marketing behaviors by funds. Further discussions will explore the relationship between fund flow sensitivity and sector concentration in the next section.

3.4 Flow Sensitivity

To demonstrate that market concentration impacts or directly relates to diseconomies of scale, I focus on its interaction with fund characteristics, particularly fund flows, which are crucial for increasing fund size. Panel A (B) of Table 3.8 presents the simple average of these characteristics in each sector for all funds (the largest three funds). Using the simple average rather than the value weighted ones eliminates the size effects.

For all funds in each sector, there is no clear correlation between HHI and other characteristics. High-concentration sectors do not necessarily have large inflows. Fund ages are around 10 years except for the two new sectors. In highly concentrated sectors, turnover rates could be either low (e.g., EDCL for large capital equity funds) or high (e.g., EDYH for hedging style investment funds). The load, fee, and expense ratio show no correlation with concentration. Some expense ratios are negative due to waivers and reimbursements. Regarding stock holding characteristics, funds would hold stocks consistent with their objectives. This implies that these characteristics are relevant to return and size but are irrelevant to market concentration. To summarize, I find none of the sector average characteristics shown here significantly correlated with the sector concentration.

Panel B displays the corresponding characteristics for the largest three funds in each sector to ensure whether the relatively large funds in each sector differ in characteristics according to sector concentration. The results are consistent, except that some large funds in highly concentrated sectors have relatively low flows but high marketing fees and expense ratios. The large funds have similar tracking errors and holding characteristics with the sector mean. This represents they are doing similar things with the other funds. However, the poor performance of large funds in concentrated sectors could be attributed to other tiny things. For example, the large funds in EDCI and EDYS have a clear large actual 12b(1) expense compared with other large funds in the other sectors.[5] However, these large funds have a similar size to other large funds in the large sectors. This aligns with our hypothesis that flow sensitivity to certain factors may differ based on the magnitude of sector concentration.

Two hypotheses are proposed:

- **Hypothesis 1**: In more concentrated sectors, flows are less sensitive to past returns. Consequently, although large funds perform poorly, they still attract inflows.
- **Hypothesis 2**: Flows are only sensitive to **positive** past returns when the concentration level is high and marketing expenses are also high. Simultaneously, flows do not react aggressively to **poor** past performance in concentrated sectors.

According to the first hypothesis, large funds would continue to attract inflows in concentrated sectors or at least prevent outflows, even if they exhibit diseconomies of scale. The second hypothesis attempts to explain this by adding the marketing factor. Funds that allocate significant resources to marketing may benefit more from previous good performance, especially in concentrated sectors. Similarly, funds that invest heavily in marketing and perform poorly in the past may not suffer from outflows or low inflows in highly concentrated sectors. This implies that large funds could employ marketing strategies to attract inflows without the risk of being penalized for poor performance.

In concentrated markets, investors tend to rely more on marketing information and give greater weight to previous positive returns when evaluating fund managers' skills. In other words, when funds are concentrated, marketing becomes a profitable strategy, and funds may be more willing to allocate resources to marketing than funds in other sectors. Consider a large fund in a concentrated sector and a fund in a distributed sector with exactly the same characteristics and holdings.

[5] The t-stat is 2.41.

Investors could compare the second fund with other funds of similar size and determine whether its marketing information is unbiased. However, for the first fund, investors have no comparable funds within the same objective category due to its outstanding size. Investors with limited attention may trust the first fund more. Consequently, the first fund is more likely to allocate more resources to marketing and may even reduce the importance of skill in portfolio management.

In a hypothetical scenario featuring a large fund in a concentrated sector alongside a fund in a distributed sector with identical characteristics and holdings, investors have different frames of reference. They can readily compare the second fund with other funds of similar size, enabling them to evaluate the accuracy of its marketing information. However, for the first fund, which stands out in terms of size within the same objective funds, investors lack comparable alternatives. With limited attention spans, investors may inherently trust the first fund, assuming its prominence reflects superior qualities. Consequently, the first fund may prioritize investing more in marketing efforts, potentially diminishing the perceived importance of portfolio management skills.

3.4.1 Past Return sensitivity

Table 3.9 presents the results of how fund flows respond to past returns at different levels of market concentration. Panel A illustrates the reaction of short-term flows to short-term and long-term past performance. Short-term flows respond positively to both short and long-term past performance, indicating a tendency to chase good performance regardless of the time horizon. However, the interaction terms with concentration exhibit significant negative coefficients in columns 1, 3, and 4. This aligns with the first hypothesis discussed earlier – flows are less sensitive to past returns in concentrated sectors. A one percentage point increase in HHI leads to a decrease in flow sensitivity to short and long-term past returns by one-tenth. Moreover, large funds tend to attract relatively smaller inflows. Although the interaction regression in column 5 is insignificant, the effect is diminished in concentrated sectors because the interaction term of size and HHI is positively significant in column 1. These findings provide evidence for Hypothesis 1.

In summary, fund flows are positively correlated with past returns and negatively correlated with fund size. However, in concentrated sectors, flows are less sensitive to past returns, whether in the short or long run. Additionally, large funds tend to attract more inflows compared to funds of the same size in less concentrated sectors. It is evident that large funds in concentrated sectors continue to attract inflows, regardless of their performance.

The results are quite similar for long-term flow in Panel B. The only difference is short-term returns play the same role at different levels of concentration in predicting long-term flow, as shown in columns 1 and 3. The magnitude of long-term flow sensitivity becomes one-tenth of that observed in short-term flow regressions, but the impact of sector concentration on the sensitivity remains unchanged.

3.4.2 Market Fee sensitivity

Since flow in concentrated sectors is not sensitive to return, what motivates investors to invest in them? According to **Hypothesis 2**, one possible reason is that, when combined with marketing expenses, flows are more sensitive to positive past returns and at least not sensitive to negative returns in more concentrated sectors. So, in Table 3.10, the Actual 12b(1) fee and its interaction terms with the positive and negative parts of past returns are added to the regression. Due to the issue that a triple interaction term would cause non-negligible multicollinearity, I sort all sectors into five groups based on their concentration level (HHI) to test Hypothesis 2. By conducting these group-by-group regressions, the monotonicity of flow sensitivity to positive and negative parts of past returns has also been tested. In these regressions, fund-month observations with null Actual 12b(1) fillings in the previous fiscal year are dropped. In 1980, the SEC approved Rule 12b(1), allowing mutual funds to deduct a certain amount of money from net assets to distribute to selling agents. The reported 12b(1) fee ratio is then a good proxy for how much a fund pays to market itself.

In the regression of Table 3.10, the quintile rank of the Actual 12b(1) ratio in the most recent fiscal year of month t is added. Its interaction terms with $Ret^{+}_{t-12} = \mathbb{1}_{Ret_{t-12}>0} Ret_{t-12}$ and $Ret^{-}_{t-12} = \mathbb{1}_{Ret_{t-12}<0} Ret_{t-12}$, where $\mathbb{1}_{Ret_{t-12}>0}$ and $\mathbb{1}_{Ret_{t-12}<0}$

are indicator functions, are also added. The coefficients of the interaction terms therefore represent how marketing efforts benefit flow sensitivity to past returns. It is true that 12b(1) is an endogenous fund choice that also relates to its size and expectations of future performance. However, these reduced form regressions in Table 3.10 suggest that flows in highly concentrated sectors are less sensitive to positive past returns if the fund has lower or no marketing expenses, as indicated in the first two rows. Funds are more sensitive to positive returns when marketing efforts are high.

The coefficients of Ret^{+}_{t-12} are similar in the first three groups, while they dramatically and monotonically drop for the last two concentrated groups. Except for the middle group, all groups are not sensitive to poor past returns, regardless of whether the marketing expense is zero (which is the second row for Ret^{-}_{t-12}) or the marketing expense is high (which is the 5th row for $Ret^{-}_{t-12} \times Actual12b(1)$). For the 12b(1) itself, the coefficients are all negative or insignificant in all groups. For the positive feedback, $Ret^{+}_{t-12} \times Actual12b(1)$, it is insignificant and even negative for less concentrated sectors, which means marketing is not efficient in these groups, even when the funds have good performance. However, in the concentrated two groups, marketing has a positive and efficient impact on past good performance. The coefficients increase from -0.015 to 0.02 then to 0.037.

Combining these results with the results in the previous section, what leads to low flow sensitivity to returns in concentrated sectors is those funds that spend less on marketing, in which flows under-react to positive past returns. However, flow sensitivity to positive returns among funds with high marketing expenses in these sectors is high. Table 3.8 shows that although the Actual 12b(1)s are similar across fund sectors, large funds in highly concentrated sectors tend to have high Actual 12b(1)s. Altogether, it is possible that large funds in concentrated sectors spend more on marketing and thus their flows are more sensitive to positive past returns but less sensitive to negative past returns. This implies they may gain flows easily from good performance without the penalty of outflow for bad performance. For small funds in these sectors, it is hard to get inflows even if they really performed well in the past and may have a good return in the future. The flows in less concentrated sectors are very sensitive to positive past returns, especially for positive ones, regardless of whether the marketing expense is high or not. Altogether, the magnitude of diseconomies of scale is larger in concentrated sectors.

Since marketing is beneficial for fund managers in concentrated sectors, it is not surprising that they invest more effort in it, regardless of their skill level. The question arises: why does marketing in concentrated sectors yield more rewards and less penalties? One possible explanation is that investors are information-limited in concentrated sectors. When they receive marketing information from large funds, they have no other sources to compare it with other similar funds in these sectors. Consequently, they are more inclined to trust the marketing and distribution information for these funds when they observe good performance. Similarly, when they observe poor performance, they may attribute it to systematic risk or pricing errors and even anticipate high expected returns in the future. Large funds in these sectors exploit this informational advantage to attract inflows, aiming to grow even larger and stand out in size. Such outcomes result in a mismatch between flow and expected returns. Managers who devote time and effort to portfolio management may not receive sufficient inflows, and their fund size may not necessarily be large. In less concentrated sectors, this recursive feedback loop is weaker and less significant, resulting in a lower magnitude of diseconomies of scale. Nevertheless, this loop never disappears, nor do the diseconomies of scale.

3.5 Conclusion

In this paper, I focus on the magnitude of mutual fund diseconomies of scale. I demonstrate that when a fund sector is concentrated, such magnitude is significant. This relationship is monotonic with respect to sector concentration, which I proxy using HHI in this paper. The interaction effect of sector and concentration even outweighs the size effect alone. One explanation for this is the difference in flow sensitivity to past returns and marketing expense in different sectors. In more concentrated sectors, flows are less sensitive to past returns. Consequently, even when large funds perform poorly, they can avoid outflows. Meanwhile, marketing is beneficial for funds in concentrated sectors. Marketing expenses in highly concentrated sectors would have a more positive feedback to flow if the fund performs well in the past, while it would have an insignificant feedback to flow if the fund performs poorly. Large funds in concentrated sectors take advantage of this by spending more on marketing, thereby increasing their size. Through this channel, large funds may not necessarily outperform.

Bibliography

Berk, Jonathan B, and Richard C Green. 2004. "Mutual fund flows and performance in rational markets." *Journal of political economy*, 112(6): 1269–1295.

Bhojraj, Sanjeev, Young Jun Cho, and Nir Yehuda. 2012. "Mutual fund family size and mutual fund performance: The role of regulatory changes." *Journal of Accounting Research*, 50(3): 647–684.

Chen, Joseph, Harrison Hong, Ming Huang, and Jeffrey D Kubik. 2004. "Does fund size erode mutual fund performance? The role of liquidity and organization." *American Economic Review*, 94(5): 1276–1302.

Elton, Edwin J, Martin J Gruber, and Christopher R Blake. 2012. "Does mutual fund size matter? The relationship between size and performance." *The Review of Asset Pricing Studies*, 2(1): 31–55.

Hoberg, Gerard, Nitin Kumar, and Nagpurnanand Prabhala. 2018. "Mutual fund competition, managerial skill, and alpha persistence." *The Review of Financial Studies*, 31(5): 1896–1929.

Huang, Shiyang, Xu Lu, Yang Song, and Hong Xiang. 2023. "Remeasuring Scale in Active Management." *Available at SSRN*.

Kacperczyk, Marcin, Clemens Sialm, and Lu Zheng. 2005. "On the industry concentration of actively managed equity mutual funds." *The Journal of Finance*, 60(4): 1983–2011.

Pástor, Luboš, and Robert F Stambaugh. 2012. "On the size of the active management industry." *Journal of Political Economy*, 120(4): 740–781.

Pástor, L'uboš, Robert F Stambaugh, and Lucian A Taylor. 2015. "Scale and skill in active management." *Journal of Financial Economics*, 116(1): 23–45.

Pollet, Joshua M, and Mungo Wilson. 2008. "How does size affect mutual fund behavior?" *The Journal of Finance*, 63(6): 2941–2969.

Song, Yang. 2017. "The mismatch between mutual fund scale and skill." *The Journal of Finance*.

Yan, Xuemin. 2008. "Liquidity, investment style, and the relation between fund size and fund performance." *Journal of Financial and Quantitative Analysis*, 741–767.

Zheng, Lu, Marcin Kacperczyk, and Clemens Sialm. 2007. "Industry concentration and mutual fund performance." *The Journal of Investment Management*, 5(1): 50–64.

Zhu, Min. 2018. "Informative fund size, managerial skill, and investor rationality." *Journal of Financial Economics*, 130(1): 114–134.

Table 3.1: Summary statistics

The variables are based on fund-month observations. TNA is the total net assets in billion dollars. Expense ratio and total load are taken for the previous fiscal year. Flow and turnover ratios are taken for the previous 12 months. Age is calculated in years. All variables are at the fund level. If the fund has more than one share class, the variables are the TNA weighted average of each share class.

Variables	Obs	Mean	Std.Dev.	Min	Max
TNA (Billion $)	626387	1.35	3.50	0.0063	24.90
Family TNA (Billion $)	621901	78.76	214	0.0132	1330
Expense Ratio	582905	1.10%	0.50%	0.00%	2.50%
12 Months Flow	590893	21.50%	98.80%	-68.10%	695.00%
Age	621409	10.024	7.349	1	35
Turnover Ratio	626387	68.10%	86.90%	0.00%	536.00%
Total Load	626387	1.40%	1.80%	0.00%	6.20%
Share Classes	626387	2.891	2.072	1	10
HHI	626387	3.47%	3.87%	1.35%	79.16%

Table 3.2: Comparison by fund sector

The variables are based on fund-month observations. Nine sectors are included in the sample. HHI is calculated using Eq. 3.2. All standard deviation, minimum and maximum value of TNA is calculated in each month and averaged over time. MAD is the mean absolute deviation of each sector. Tracking Error is the fund's average tracking error based on its benchmark. Spatial distance is calculated using a similar method with Hoberg, Kumar and Prabhala (2018).

CRSP Obj. Code	EDCI	EDCL	EDCM	EDCS	EDYB	EDYG	EDYH	EDYI	EDYS
Num. Month	252	252	252	252	252	252	168	252	168
HHI	16.43%	17.30%	2.65%	2.03%	4.58%	1.78%	6.04%	5.78%	41.56%
Num. Funds	34.00	58.90	287.75	444.34	491.03	947.30	151.48	114.53	10.24
TNA (Million)	310.71	10450.57	1313.72	938.61	2377.87	1603.20	396.69	1876.30	370.59
SD TNA (Million)	658.17	30244.97	3415.07	2777.15	12030.85	6263.04	1012.87	4152.79	721.76
Min TNA (Million)	8.53	31.40	5.76	5.41	5.45	5.21	5.98	5.94	8.41
Max tna (Million)	3749.54	159638.13	35639.34	38725.38	227467.45	129995.81	7983.38	26107.57	2269.28
Largest TNA %	35.30%	30.01%	8.45%	8.75%	15.16%	8.78%	14.35%	13.87%	59.17%
Top 3 TNA %	51.46%	65.51%	20.12%	18.07%	27.25%	16.82%	33.72%	32.11%	85.97%
Top 5 TNA %	60.79%	79.52%	27.91%	24.24%	35.89%	21.73%	45.55%	42.99%	94.46%
Top 10 TNA %	76.94%	87.98%	40.54%	34.63%	47.85%	31.32%	62.07%	60.87%	99.22%
MAD of TNA	408.93	18021.91	2017.03	1437.84	4100.78	2649.78	626.55	2911.60	538.79
Tracking Error	4.76%	3.34%	3.93%	4.50%	3.36%	3.63%	5.46%	3.53%	12.92%
Spatial Distance	1.06	0.01	0.71	0.70	0.69	0.91	0.99	0.64	0.12
SD TNA/Mean	2.07	3.05	2.49	2.81	4.59	3.97	2.53	2.26	1.82
MAD/Mean	1.31	1.70	1.53	1.51	1.71	1.66	1.57	1.55	1.41

Table 3.3: Sector and tracking index

CRSP Obj. Code	Tracking Index
EDCI	US Micro Cap Index
EDCL	US Large Cap Index
EDCM	US Mid Cap Index
EDCS	US Small Cap Index
EDYB EDYG	Avg of Mega, Large, Mid, Small and Micro Growth Index
EDYH	US Total Market Index
EDYI	Avg of Mega, Large, Mid, Small and Micro Income Index
EDYS	US Total Market Index

Table 3.4: Introduction of the market concentration

The dependent variables in Panel A are all based on fund net return. The **Raw** is the return itself as the dependent variable. **Ret-Mkt_Ret** uses raw return minus the market portfolio return as the dependent variable. **Ret_Beta** uses CAPM model alpha and residual, while **Ret_FF3** and **Ret_FF4** use the Fama French three-factor model and such three factors together with the momentum factor, respectively. Based on Chen et al. (2004), all factor models are estimated in fund size group level. Panel B has all the same settings except that it is based on gross return, which equals net return plus the monthly expense ratio of each fund. TNA and family TNA are in million-dollar level, and HHI has been multiplied by 10,000 resulting in the range from 0 to 10,000. All these variables are taken logarithmically. Exp_Ratio and Total Load are the last fiscal year's expense ratio and total load. Flow and turnover ratio represent the flow plus turnover from month $t-12$ to month $t-1$. Age is the fund's age in years. All regressions are at the fund-month level, using Fama-Macbeth regression method. All standard errors are based on fourth-order Newey-West adjustment. T-statistics are reported in parentheses.

(a) Net return

	Raw	Ret-Mkt_Ret	Ret_Beta	Ret_FF3	Ret_FF4
Log TNA	-0.030**	-0.030**	-0.040***	-0.039***	-0.039***
	(-2.50)	(-2.50)	(-3.34)	(-3.84)	(-3.88)
Log HHI	-0.105***	-0.105***	-0.105***	-0.105***	-0.105***
	(-2.60)	(-2.60)	(-2.60)	(-2.61)	(-2.62)
Log FamTNA	0.020***	0.020***	0.020***	0.020***	0.020***
	(3.53)	(3.53)	(3.56)	(3.51)	(3.51)
Exp_Ratio	-3.803	-3.803	-3.577	-3.584	-3.583
	(-1.00)	(-1.00)	(-0.95)	(-0.95)	(-0.95)
Flow	0.006	0.006	0.006	0.006	0.006
	(0.33)	(0.33)	(0.30)	(0.31)	(0.30)
Age	0.006***	0.006***	0.005***	0.006***	0.006***
	(2.97)	(2.97)	(2.95)	(2.98)	(2.98)
Turnover	-0.019	-0.019	-0.02	-0.02	-0.02
	(-0.50)	(-0.50)	(-0.50)	(-0.51)	(-0.51)
Total Load	-1.933***	-1.933***	-2.000***	-1.975***	-1.973***
	(-3.00)	(-3.00)	(-3.09)	(-3.04)	(-3.03)
Constant	1.244***	0.572***	0.641***	0.645***	0.646***
	(2.93)	(2.89)	(3.24)	(3.40)	(3.40)
Observations	544499	544499	544499	544499	544499

Table 3.4: Introduction of the market concentration

(b) Gross return

	Raw	Ret-Mkt_Ret	Ret_Beta	Ret_FF3	Ret_FF4
Log TNA	-0.030**	-0.030**	-0.031***	-0.031***	-0.031***
	(-2.51)	(-2.51)	(-2.65)	(-3.03)	(-3.06)
Log HHI	-0.105**	-0.105**	-0.104**	-0.105***	-0.105***
	(-2.59)	(-2.59)	(-2.59)	(-2.61)	(-2.61)
Log FamTNA	0.020***	0.020***	0.020***	0.020***	0.020***
	(3.49)	(3.49)	(3.52)	(3.48)	(3.48)
Exp_Ratio	4.828	4.828	5.049	5.041	5.042
	(1.27)	(1.27)	(1.34)	(1.34)	(1.34)
Flow	0.006	0.006	0.006	0.006	0.006
	(0.33)	(0.33)	(0.30)	(0.30)	(0.30)
Age	0.006***	0.006***	0.006***	0.006***	0.006***
	(2.97)	(2.97)	(2.96)	(2.99)	(2.99)
Turnover	-0.02	-0.02	-0.02	-0.02	-0.02
	(-0.50)	(-0.50)	(-0.51)	(-0.51)	(-0.51)
Total Load	-1.935***	-1.935***	-2.006***	-1.981***	-1.979***
	(-3.01)	(-3.01)	(-3.10)	(-3.05)	(-3.04)
Constant	1.243***	0.571***	0.512**	0.516***	0.517***
	(2.93)	(2.88)	(2.59)	(2.72)	(2.72)
Observations	544499	544499	544499	544499	544499

Table 3.5: Performance in concentrated sectors

The dependent variables in Panel A are all based on fund net return and Panel B has all the same settings except that it is based on gross return, which equals net return plus the monthly expense ratio of each fund. $D(HHI)$ is the dummy for the HHI index when the fund$i$ is in the top quintile of its category in month$t$. All other variables are identical to Table 3.4. All regressions are at the fund-month level, using the Fama-Macbeth regression method. All standard errors are based on the fourth order Newey-West adjustment. T-statistics are reported in parentheses.

(a) Net return

	Raw	Ret-Mkt_Ret	Ret_Beta	Ret_FF3	Ret_FF4
Log TNA	-0.029**	-0.029**	-0.039***	-0.039***	-0.039***
	(-2.46)	(-2.46)	(-3.30)	(-3.81)	(-3.86)
D(HHI)	-0.124**	-0.124**	-0.123**	-0.124**	-0.124**
	(-2.55)	(-2.55)	(-2.55)	(-2.59)	(-2.59)
Log FamTNA	0.020***	0.020***	0.020***	0.019***	0.019***
	(3.30)	(3.30)	(3.33)	(3.29)	(3.29)
Exp_Ratio	-3.945	-3.945	-3.704	-3.735	-3.732
	(-0.97)	(-0.97)	(-0.92)	(-0.92)	(-0.92)
Flow	0.005	0.005	0.005	0.005	0.005
	(0.28)	(0.28)	(0.25)	(0.25)	(0.25)
Age	0.006***	0.006***	0.006***	0.006***	0.006***
	(3.01)	(3.01)	(2.98)	(3.02)	(3.02)
Turnover	-0.02	-0.02	-0.02	-0.02	-0.02
	(-0.52)	(-0.52)	(-0.53)	(-0.53)	(-0.53)
Total Load	-1.825***	-1.825***	-1.894***	-1.866***	-1.864***
	(-2.84)	(-2.84)	(-2.93)	(-2.88)	(-2.87)
Constant	0.671***	-0.001	0.071	0.072	0.072
	(2.72)	(-0.00)	(0.63)	(0.67)	(0.68)
Observations	544499	544499	544499	544499	544499

Table 3.5: Performance in concentrated sectors

(b) Gross return

	Raw	Ret-Mkt_Ret	Ret_Beta	Ret_FF3	Ret_FF4
Log TNA	-0.029**	-0.029**	-0.031***	-0.030***	-0.030***
	(-2.47)	(-2.47)	(-2.62)	(-3.01)	(-3.04)
D(HHI)	-0.123**	-0.123**	-0.123**	-0.124**	-0.124**
	(-2.54)	(-2.54)	(-2.54)	(-2.58)	(-2.58)
Log FamTNA	0.019***	0.019***	0.019***	0.019***	0.019***
	(3.27)	(3.27)	(3.29)	(3.25)	(3.25)
Exp_Ratio	4.686	4.686	4.921	4.891	4.894
	(1.15)	(1.15)	(1.22)	(1.21)	(1.21)
Flow	0.005	0.005	0.005	0.005	0.005
	(0.27)	(0.27)	(0.25)	(0.25)	(0.25)
Age	0.006***	0.006***	0.006***	0.006***	0.006***
	(3.01)	(3.01)	(2.99)	(3.03)	(3.03)
Turnover	-0.02	-0.02	-0.021	-0.021	-0.021
	(-0.52)	(-0.52)	(-0.53)	(-0.53)	(-0.53)
Total Load	-1.827***	-1.827***	-1.900***	-1.872***	-1.870***
	(-2.84)	(-2.84)	(-2.94)	(-2.88)	(-2.88)
Constant	0.671***	-0.001	-0.057	-0.056	-0.056
	(2.72)	(-0.01)	(-0.51)	(-0.53)	(-0.53)
Observations	544499	544499	544499	544499	544499

Table 3.6: Interaction between the concentration and the size

This table reports the regression results identical to the model in Table 3.4. The only difference is we introduce the log HHI index or HHI ranking and its interaction with fund size in Panel A or B. All regressions are at the fund-month level, using the Fama-Macbeth regression method. All standard errors are based on the fourth order Newey-West adjustment. T-statistics are reported in parentheses.

(a) Log HHI as concentration

	Raw	Ret-Mkt_Ret	Ret_Beta	Ret_FF3	Ret_FF4
Log TNA	0.089	0.089	0.078	0.077	0.078
	(1.48)	(1.48)	(1.30)	(1.29)	(1.29)
Log HHI× Log TNA	-0.021**	-0.021**	-0.021**	-0.021**	-0.021**
	(-2.01)	(-2.01)	(-1.99)	(-1.99)	(-1.99)
Quintile Rank of HHI	0.049	0.049	0.047	0.044	0.044
	(0.52)	(0.52)	(0.50)	(0.48)	(0.48)
Log FamTNA	0.018***	0.018***	0.019***	0.018***	0.018***
	(3.38)	(3.38)	(3.41)	(3.36)	(3.37)
Exp_Ratio	-5.268	-5.268	-4.994	-5.026	-5.026
	(-1.50)	(-1.50)	(-1.43)	(-1.44)	(-1.44)
Flow	0.006	0.006	0.006	0.006	0.006
	(0.32)	(0.32)	(0.29)	(0.30)	(0.29)
Age	0.006***	0.006***	0.005***	0.006***	0.006***
	(2.96)	(2.96)	(2.94)	(2.98)	(2.98)
Turnover	-0.021	-0.021	-0.021	-0.021	-0.021
	(-0.53)	(-0.53)	(-0.54)	(-0.54)	(-0.54)
Total Load	-1.727***	-1.727***	-1.800***	-1.770***	-1.768***
	(-2.93)	(-2.93)	(-3.03)	(-2.98)	(-2.97)
Constant	0.667***	-0.005	0.066	0.067	0.068
	(2.61)	(-0.04)	(0.59)	(0.64)	(0.65)
Observations	544499	544499	544499	544499	544499

Table 3.6: Interaction between the concentration and the size

(b) HHI ranking as concentration

	Raw	Ret-Mkt_Ret	Ret_Beta	Ret_FF3	Ret_FF4
Log TNA	-0.01	-0.01	-0.02	-0.02	-0.02
	(-0.78)	(-0.78)	(-1.55)	(-1.58)	(-1.61)
Quintile Rank of HHI×Log TNA	-0.015**	-0.015**	-0.015**	-0.014**	-0.014**
	(-2.10)	(-2.10)	(-2.09)	(-2.09)	(-2.09)
Quintile Rank of HHI	0.071	0.071	0.071	0.067	0.067
	(0.71)	(0.71)	(0.70)	(0.67)	(0.67)
Log FamTNA	0.019***	0.019***	0.019***	0.018***	0.018***
	(3.40)	(3.40)	(3.43)	(3.39)	(3.39)
Exp_Ratio	-5.014	-5.014	-4.754	-4.784	-4.781
	(-1.37)	(-1.37)	(-1.31)	(-1.32)	(-1.32)
Flow	0.008	0.008	0.007	0.007	0.007
	(0.41)	(0.41)	(0.38)	(0.38)	(0.38)
Age	0.005***	0.005***	0.005***	0.005***	0.005***
	(2.92)	(2.92)	(2.89)	(2.93)	(2.93)
Turnover	-0.019	-0.019	-0.019	-0.019	-0.019
	(-0.49)	(-0.49)	(-0.49)	(-0.49)	(-0.50)
Total Load	-1.718***	-1.718***	-1.789***	-1.760***	-1.758***
	(-2.82)	(-2.82)	(-2.92)	(-2.86)	(-2.86)
Constant	0.660***	-0.012	0.059	0.06	0.06
	(2.60)	(-0.11)	(0.52)	(0.57)	(0.58)
Observations	544499	544499	544499	544499	544499

Table 3.7: Size effect in sectors with different levels of concentration

This table reports the estimation of regressions identical to Table 3.4 but with subsamples. Each column represents a subsample regression that is highlighted at the column head. The HHI index are sorted in each fund sector into five groups with a higher quintile representing a higher HHI value. The groups are rebalanced each month. All regressions are at the fund-month level, using the Fama-Macbeth regression method. All standard errors are based on the fourth order Newey-West adjustment. T-statistics are reported in parentheses.

(a) Net return

Quintile of HHI	1	2	3	4	5
Log TNA	-0.023	-0.037***	-0.052**	-0.024	-0.113**
	(-1.29)	(-3.32)	(-2.21)	(-0.76)	(-2.17)
Log FamTNA	0.009*	0.017**	0.021*	0.018	0.001
	(1.66)	(2.26)	(1.66)	(0.89)	(0.04)
Exp_Ratio	-12.152***	-8.620**	-6.69	14.441	-22.049
	(-3.29)	(-2.16)	(-0.90)	(1.34)	(-1.13)
Flow	0.03	0.008	0.022	0.064*	-0.343
	(1.43)	(0.45)	(0.82)	(1.90)	(-1.58)
Age	0.001	0.002	0.013***	0.005	0.045***
	(0.67)	(1.32)	(2.92)	(1.33)	(3.21)
Turnover	-0.032	-0.024	-0.02	0.039	0.517***
	(-0.54)	(-0.75)	(-0.50)	(0.72)	(3.34)
Total Load	-0.136	-0.221	-2.451*	-7.726**	16.770***
	(-0.24)	(-0.40)	(-1.70)	(-2.01)	(3.58)
Constant	0.353*	0.147	0.104	-0.304	-0.730**
	(1.76)	(1.12)	(0.50)	(-1.53)	(-2.08)
Observations	155848	155848	155848	155848	155848

Table 3.7: Size effect in sectors with different levels of concentration

(b) Gross return

Quintile of HHI	1	2	3	4	5
Log TNA	-0.015	-0.029**	-0.044*	-0.017	-0.105**
	(-0.84)	(-2.58)	(-1.87)	(-0.52)	(-2.01)
Log FamTNA	0.009	0.017**	0.021*	0.018	0.000
	(1.61)	(2.24)	(1.65)	(0.89)	0.00
Exp_Ratio	-3.353	0.051	1.797	22.870**	-13.401
	(-0.91)	(0.01)	(0.24)	(2.12)	(-0.69)
Flow	0.03	0.008	0.022	0.064*	-0.343
	(1.42)	(0.44)	(0.82)	(1.92)	(-1.58)
Age	0.001	0.002	0.013***	0.005	0.044***
	(0.69)	(1.34)	(2.92)	(1.33)	(3.21)
Turnover	-0.032	-0.024	-0.02	0.04	0.518***
	(-0.54)	(-0.75)	(-0.51)	(0.72)	(3.36)
Total Load	-0.147	-0.246	-2.454*	-7.703**	16.703***
	(-0.26)	(-0.44)	(-1.70)	(-2.01)	(3.56)
Constant	0.224	0.017	-0.023	-0.431**	-0.853**
	(1.12)	(0.13)	(-0.11)	(-2.18)	(-2.43)
Observations	155848	155848	155848	155848	155848

Table 3.8: Fund characteristics in each sector

All Observations in this table are at the fund-month level. Share classes, flow, age, turnover, total load, and tracking error have the same definitions as in Table 3.1 and Table 3.2. Average stock coverage represents how many analysts cover each holding stock in the previous 24 months. Number of stocks is the number of different stocks held by each fund. Log size, BM ratio, MOM are the average stock characteristics of the fund's holdings. The corresponding next three rows are the NYST percentile of these variables. Actual 12b(1) is the marketing and distribution fee for such month in this fiscal year. Management fee is the total management fee divided by average net assets. Expense ratio is the ratio of the total investment that shareholders pay for the fund's operating expenses, which include 12b(1) fees.

(a) All funds

CRSP Obj. Code	EDCI	EDCL	EDCM	EDCS	EDYB	EDYG	EDYH	EDYI	EDYS
HHI	16.43%	17.30%	2.65%	2.03%	4.58%	1.78%	6.04%	5.78%	41.56%
Num. Month	252	252	252	252	252	252	168	252	168
Avg. Share Classes	1.85	2.69	3.09	2.84	3.17	2.80	2.47	2.96	1.76
12 Month Flow	13.28%	8.68%	22.85%	14.99%	19.28%	22.90%	46.24%	29.35%	51.59%
Age	11.57	11.90	10.16	10.43	9.60	10.04	6.56	10.12	5.55
Turnover	73.67%	9.07%	80.11%	77.54%	51.53%	76.64%	222.36%	46.56%	153.61%
Total Load	1.43%	0.66%	1.40%	1.34%	1.54%	1.48%	0.93%	1.53%	0.17%
Avg. Stock Analyst Coverage	54.16	29.38	33.79	31.32	32.80	49.46	49.49	29.29	34.40
Num.Stocks	32.59	368.91	95.64	92.81	84.60	83.75	128.97	64.13	80.63
Log Size	12.89	17.77	15.41	14.22	17.22	16.92	16.58	17.21	17.36
BM Ratio	0.74	0.44	0.49	0.58	0.49	0.44	0.55	0.50	0.59
MOM	38.04%	14.52%	24.57%	26.10%	15.18%	20.99%	18.51%	12.61%	15.30%
NYSE Percentile of Size	19.66%	94.04%	69.68%	45.55%	88.71%	85.51%	78.54%	88.31%	89.55%
NYSE Percentile of BM	53.24%	34.40%	37.83%	44.74%	38.96%	33.36%	43.52%	39.78%	38.87%
NYSE Percentile of MOM	60.74%	54.24%	59.16%	58.93%	54.57%	57.56%	56.36%	53.16%	52.35%
Actual 12b(1)	0.44%	0.42%	0.51%	0.51%	0.53%	0.53%	0.55%	0.53%	0.37%
Management Fee	0.71%	0.08%	0.46%	0.21%	0.05%	0.29%	0.69%	0.52%	0.67%
Expense Ratio	1.52%	-3.68%	-2.48%	-7.67%	-2.56%	-7.68%	1.61%	-0.27%	1.33%
Tracking Error	4.76%	3.34%	3.93%	4.50%	3.36%	3.63%	5.46%	3.53%	12.92%

(b) Largest 3 funds in each sector

CRSP Obj. Code	EDCI	EDCL	EDCM	EDCS	EDYB	EDYG	EDYH	EDYI	EDYS
HHI	16.43%	17.30%	2.65%	2.03%	4.58%	1.78%	6.04%	5.78%	41.56%
Num. Month	252	252	252	252	252	252	168	252	168
Avg. Share Classes	1.91	1.88	3.32	2.92	8.66	5.23	3.57	3.24	2.25
12 Month Flow	-3.75%	6.34%	3.60%	2.43%	-0.38%	-3.47%	0.93%	-0.85%	43.73%
Age	14.37	15.27	15.28	15.92	9.94	20.65	9.42	17.72	6.08
Turnover	48.85%	5.31%	46.85%	26.97%	21.59%	36.13%	129.03%	35.27%	239.33%
Total Load	1.23%	0.03%	0.76%	0.95%	2.55%	1.47%	0.97%	1.14%	0.23%
Avg. Stock Analyst Coverage	48.15	29.43	33.12	29.40	31.18	41.18	33.13	29.30	35.89
Num.Stocks	79.31	370.37	286.75	322.31	336.93	130.60	141.65	85.95	53.30
Log Size	12.95	17.77	15.62	14.50	17.55	17.36	17.29	17.46	17.67
BM Ratio	0.71	0.44	0.46	0.53	0.49	0.38	0.51	0.47	0.67
MOM	43.18%	14.58%	28.24%	26.82%	13.86%	20.99%	16.84%	12.18%	15.61%
NYSE Percentile of Size	20.78%	94.05%	71.95%	51.58%	92.94%	91.53%	88.20%	92.00%	92.37%
NYSE Percentile of BM	51.63%	34.32%	35.50%	40.55%	38.21%	28.72%	40.42%	37.83%	42.87%
NYSE Percentile of MOM	58.78%	54.12%	59.65%	58.98%	53.86%	58.44%	55.63%	52.78%	52.10%
Actual 12b(1)	0.57%	0.02%	0.31%	0.27%	0.42%	0.47%	0.54%	0.40%	0.51%
Management Fee	0.95%	0.05%	0.36%	0.47%	0.27%	0.46%	0.92%	0.44%	0.75%
Expense Ratio	1.23%	0.09%	0.64%	0.68%	0.66%	0.78%	1.37%	0.78%	1.22%
Tracking Error	4.71%	3.31%	3.83%	4.28%	3.40%	3.58%	3.30%	3.42%	12.74%

Table 3.9: Flow sensitivity to the past performance

This table reports the regression coefficients for a panel regression as follows,

$$y_{i,t} = \beta_1 R_{i,t-1} + \beta_2 Ret_{i,t-1} Log(HHI_{i,t-1}) + \beta_3 Ret_{t-12} + \beta_4 Ret_{i,t-2,t-13} Log(HHI_{i,t-1}) + \textbf{Controls}.$$

In Panel A, the dependent variable is $Flow_{i,t,t+1}$, while in Panel B, the dependent variable is $Flow_{i,t+1,t+12}$. Ret_{t-1} represents the monthly net raw return in month $t-1$, while $Ret_{t-2,t-13}$ is the net raw return from $t-12$ to $t-2$. $Log(HHI)$ is the logarithm of $(HHI \times 10000)$ in month $t-1$, and $Log(TNA)$ is also the logarithm of TNA in million dollars in month $t-1$. "Quintile Rank of HHI" is rebalanced every month and sorts the HHI of nine sectors into five groups. It is also the value of month $t-1$. The control variables are shown in the tale. All regressions are at the fund-month level, using the Fama-Macbeth regression method. All standard errors are based on fourth-order Newey-West adjustment. T-statistics are reported in parentheses.

(a) Flow from month t to $t+1$

	(1)	(2)	(3)	(4)	(5)
Ret_{t-1}	0.291***	0.149***	0.333***	0.154***	0.150***
	(3.93)	(9.80)	(5.00)	(9.89)	(9.81)
$Ret_{t-1} \times Log(HHI)$	-2.04		-2.897**		
	(-1.50)		(-2.46)		
Ret_{t-12}	0.216***	0.055***	0.056***	0.175***	0.056***
	(6.96)	(9.91)	(10.16)	(6.59)	(9.92)
$Ret_{t-12} \times Log(HHI)$	-2.647***			-1.984***	
	(-5.31)			(-4.71)	
$Log(TNA)$	-0.386***	-0.131***	-0.131***	-0.133***	-0.228***
	(-6.19)	(-18.28)	(-18.15)	(-18.01)	(-5.77)
$Log(TNA) \times Log(HHI)$	0.045***				0.017**
	(3.90)				(2.35)
Quintile Rank of HHI	-0.019	0.110**	0.191***	0.200**	-0.037
	(-0.19)	(2.41)	(2.61)	(2.58)	(-0.40)
$Flow_{t-12,t-1}$	1.369***	1.385***	1.383***	1.374***	1.384***
	(32.28)	(32.77)	(32.85)	(32.39)	(32.68)
Constant	0.325**	0.457***	0.399***	0.385**	0.457***
	(1.99)	(2.93)	(2.60)	(2.37)	(2.94)
Observations	590893	590893	590893	590893	590893

Table 3.9: Flow sensitivity to the past performance

(b) Flow from month $t+1$ to $t+12$

	(1)	(2)	(3)	(4)	(5)
Ret_{t-1}	0.023**	0.012***	0.021**	0.013***	0.012***
	(2.08)	(4.89)	(2.26)	(5.22)	(5.02)
$Ret_{t-1} \times Log(HHI)$	-0.159		-0.132		
	(-0.78)		(-0.76)		
$Ret_{t-2,t-12}$	0.026***	0.008***	0.008***	0.021***	0.008***
	(5.10)	(6.75)	(6.88)	(5.03)	(6.89)
$Ret_{t-2,t-12} \times Log(HHI)$	-0.309***			-0.232***	
	(-3.44)			(-3.23)	
$Log(TNA)$	-0.111***	-0.078***	-0.078***	-0.079***	-0.087***
	(-6.66)	(-16.27)	(-16.35)	(-16.34)	(-7.28)
$Log(TNA) \times Log(HHI)$	0.006**				0.002
	(2.03)				(0.75)
Quintile Rank of HHI	0.019	0.018	0.036*	0.025	0.006
	(0.63)	(1.50)	(1.87)	(1.25)	(0.22)
$Flow_{t-12,t-1}$	0.182***	0.184***	0.184***	0.183***	0.184***
	(21.03)	(20.84)	(21.02)	(20.92)	(20.86)
Constant	0.511***	0.534***	0.529***	0.523***	0.530***
	(11.94)	(12.78)	(12.60)	(12.18)	(12.74)
Observations	546588	546588	546588	546588	546588

Table 3.10: Flow sensitivity to past returns and marketing expenses

This table presents regression results similar to those in Model 3.9, but with observations from subsamples. Nine sectors are categorized into five groups based on the month $t-1$ HHI for month t. Each column represents a regression for all funds in such quintile group. The dependent variables are $Flow_{i,t+1,t+12}$. Ret_{t-12} denotes the net raw return from $t-12$ to $t-1$. Ret^{+}_{t-12} (Ret^{-}_{t-12}) equals Ret_{t-12} if $Ret_{t-12} > 0$ ($Ret_{t-12} < 0$), and zero otherwise. Actual 12b(1) is the marketing and distribution cost of the last fiscal year. Log TNA represents the logarithm of TNA in million dollars in month $t-1$. All regressions are conducted at the fund-month level, employing the Fama-Macbeth regression method. Standard errors are adjusted using a fourth-order Newey-West procedure. T-statistics are reported in parentheses.

Quintile of HHI	1	2	3	4	5
Ret^{+}_{t-12}	0.022***	0.066*	0.038**	0.019	0.019
	(4.01)	(1.79)	(2.34)	(0.85)	(1.39)
Ret^{-}_{t-12}	-0.001	0.044	0.016**	0.059	-0.051
	(-0.04)	(1.47)	(2.26)	(0.81)	(-1.00)
$Actual12b(1)$	-0.237*	-0.120***	-0.033	-0.229*	0.171
	(-1.66)	(-3.39)	(-0.45)	(-1.66)	(0.82)
$Ret^{+}_{t-12} \times Actual12b(1)$	0.000	-0.013	-0.015*	0.020**	0.037**
	(0.23)	(-1.33)	(-1.76)	(2.16)	(1.98)
$Ret^{-}_{t-12} \times Actual12b(1)$	-0.004	-0.039*	0.001	0.004	0.077
	(-1.21)	(-1.68)	(0.42)	(0.16)	(1.56)
$Log(TNA)$	-0.116***	-0.076***	-0.082***	-0.131***	-0.01
	(-7.48)	(-14.61)	(-10.81)	(-5.31)	(-0.49)
$Flow_{t-12,t-1}$	0.209***	0.170***	0.209***	0.250***	0.183***
	(9.86)	(15.24)	(7.56)	(5.86)	(2.66)
Constant	0.707***	0.444***	0.510***	1.008***	-0.215
	(6.34)	(6.87)	(6.04)	(5.97)	(-1.00)
Observations	93893	143383	62032	34746	6291