# Barking Up The Wrong Tree: Return-chasing in 401(k) Plans

Anh Tran* Pingle Wang†

November 15, 2022

This paper examines investors' retirement savings allocation using a hand-collected dataset on 401(k) plans. We find that 83% of investors in our sample hold only 39% of total assets and follow a return-chasing strategy. In contrast, the remaining 17% of wealthy investors with relatively higher financial literacy follow CAPM alpha. This difference between the two investor groups explains why fund flows respond to returns at the plan level but to CAPM alpha at the aggregated fund level. Return-chasing by unwealthy investors is not optimal, as it significantly underperforms a strategy that passively invests in the existing funds in their plans.

***Keywords:*** Mutual funds, 401(k), fund flows, wealth inequality, financial literacy

*Anh Tran, Department of Finance, University of Connecticut, email: anh.2.tran@uconn.edu.

†Pingle Wang, Jindal School of Management, University of Texas at Dallas, email: pingle.wang@utdallas.edu. The authors thank Xing Huang, Yang Sun, anonymous referees, and participants at University of Connecticut, University of Rochester, FMA meeting, and MFA meeting for helpful discussions and suggestions.

# 1 Introduction

How investors allocate their wealth is one of the most important questions in asset management literature. Recent studies suggest that investors value funds that deliver superior returns in excess of market exposure (Barber, Huang, and Odean (2016), Berk and van Binsbergen (2016)). Evans and Sun (2018) and Ben-David, Li, Rossi, and Song (2019) argue that observed flow-performance sensitivity is driven by Morningstar ratings, and investors rely on third-party ratings to make investment decisions. Since these studies only observed money flows at the fund level, their results characterize the representative agent's investment decision in the economy. Understanding the decision-making of the representative agent is crucially important, but it is also important to study how individual investors make their investment decisions.

Employees who contribute to 401(k) plans represent an ideal population to investigate individuals' investment decision making. In contrast to the literature that focuses on flow-performance relationship at the aggregated fund level, which is mainly driven by large institutions and wealthy individuals, our paper is the first to test which performance metrics individual investors respond to at the plan level and to examine whether there is any heterogeneity in investment behavior across 401(k) plans. One key advantage of using 401(k) plans is that they have a contribution limit, which weakens the role of wealthy individuals in determining the flow-performance relationship within a plan. We hand-collected a large sample of 401(k) plans for over 1,500 US public firms between 1993 and 2016 from annual Form 11-K filings. The granularity of the data allows us not only to study how individuals make investment decisions among mutual funds from 401(k) plans, but also to examine how they choose between their own company stock and funds.

Over the years, 401(k) plans have grown in popularity, becoming a large part of the market. Over 55 million American workers are active participants in the plan, with over \$4.7 trillion having been invested at the end of 2016.[1] Moreover, according to the recent Foundation's National Financial Capability Study, only 20% of households invest outside of a retirement account. There is, however, a lack of evidence regarding the pattern and quality of household investment decisions. This is disconcerting, given that households rely heavily on their 401(k) to generate retirement income, highlighting the importance of understanding the investment behavior of households. To understand this issue, our dataset serves as an excellent resource for researchers and policymakers.

Our analysis of 401(k) plans yields several important findings. First, we show that, at the plan level, flows are responsive to past returns rather than risk-adjusted alphas or third-party ratings. This pattern exists not only when employees choose among mutual funds, but also when they select between the company stock and funds. Moreover, flows respond to returns in the most recent year rather than prior years. Interestingly, when we aggregate fund flows across 401(k) plans, we find that flows respond to the CAPM alpha at the aggregated fund level, consistent with findings in the existing literature. We show that the seemingly different results between plan-level and aggregated fund-level can be explained by heterogeneous flow-performance relations across wealth levels. In particular, over 61% of the aggregated wealth in our 401(k) plans is held by only 17% of investors, who are relatively more financially sophisticated and invest their money following the CAPM alpha, whereas the remaining 83% of investors follow returns. Lastly, we show that, a simple strategy that passively investing in the existing funds of the plan significantly outperforms the return-chasing strategy of these unsophisticated investors.

To explain these results, we first show that investors choose mutual funds in 401(k)

[1] See Investment Company Institute (https://www.ici.org/policy/retirement/plan/401k/faqs_401k).

plans based mostly on net-of-fee returns at the plan level. We begin our flow-performance analysis by estimating horse-race panel regressions of fund flows on various performance measures, such as fund's net-of-fee returns, CAPM alpha, four-factor alpha, and Morningstar standardized returns. The regressions are similar to the ones in Barber et al. (2016) but are estimated at the plan level with a plan×year fixed effect. Our triplet panel setting at the plan-fund-time level allows us to estimate the within-plan flow-performance relation and investigate the allocation decision of the majority of investors. We find that, in the horse-race, net-of-fee returns win hands down in determining flows. To confidently establish the flow-return relation, we run two additional tests that are less parametric than the linear regression model, alleviating the concern that the convex flow-performance relation cannot be addressed in linear regression models. In the first test, following Barber et al. (2016), we exploit within-plan variations in performance rankings resulting from different performance metrics and run pairwise comparisons among all performance metrics. In the second test, following Ben-David et al. (2019), we rank funds within the plan into terciles based on each performance measure and test which measure can explain the most variations in fund flows. Both tests show that flows are mostly explained by fund returns. Our finding demonstrates that, even though the representative agent cares about the CAPM alpha or Morningstar ratings as documented in prior studies, most individual investors in our sample simply follow funds' past returns.

One potential explanation for the apparent contradiction between our findings and those of prior studies could be that investors in our sample are drastically different from those in the other studies. To rule out the sample selection bias, we aggregate our data at the fund level and re-run horse-race regressions of flows on alternative performance metrics. Interestingly, we find that the CAPM alpha and Morningstar ratings can explain future flows at the aggregated fund level, consistent with findings in the existing literature; returns, however, do not explain fund-level flows. Another possible explanation for the contradiction could be that menu change in 401(k) can induce trading and affect flow-performance relation. For example, Sialm, Starks, and Zhang (2015) find that adding or deleting funds in 401(k) plans can significantly affect fund flows. Pool, Sialm, and Stefanescu (2016) find that firms tend to add (delete) funds that outperform (underperform) in the past. To tease out the effect of menu change, we decompose fund flows into sponsor flows and participant flows following Sialm et al. (2015) and re-examine the flow-performance relation both at the fund level and plan level. We show that sponsor flows respond to CAPM alpha and Morningstar ratings both at the fund level and plan level. Participant flows, on the other hand, respond to CAPM alpha at the fund level but simply follow returns at the plan level, regardless of whether the plan menu is changed.

The seemingly contradictory evidence at the plan level and aggregated fund level for individual investors can be explained by the cross-sectional variations in wealth accumulation and investment strategies. We split our sample according to the average savings per employee into four groups and study the flow-performance relation within each group. We find that employees in the highest-savings group (hereafter, *wealthy group*) allocate their savings into funds based on the CAPM alpha, whereas employees in the remaining groups (hereafter, *unwealthy group*) chase fund returns. More importantly, the dispersion in average savings is quite large. Specifically, although the wealthy group constitutes only 17% of all employees in our sample, they hold more than 61% of the savings. They save on average $121,616, which is 27 (4.4) times greater than those in the lowest-savings (second-highest-savings) group. Therefore, the flow-performance relation tilts towards the CAPM alpha at aggregated fund level, which is mostly driven by the wealthy group and is similar to the existing findings in the literature. In contrast, the relation at the plan level is dominated by the majority of investors favoring returns.

Our finding raises the question of what forces are driving the savings gap, and why the wealthy group follows the CAPM alpha, while the unwealthy group follows net-of-fee returns. To answer these questions, we examine key determinants of savings in the 401(k) plan, including firms' matching contribution policy, plan performance, employees' job tenure, wages, and deferral rate.[2] We show that the first three determinants jointly explain only 25% of the savings gap, leaving a large proportion to be explained by wages and deferral rates.

We argue that financial literacy affects wages and deferral rates, leading to a massive difference in overall savings. Previous literature has provided empirical evidence that financial literacy, human capital, and stock market participation are endogenously determined (Lusardi and Mitchell (2014), and Spataro and Corsini (2017)). The gap in financial literacy between wealthy and unwealthy groups can also explain why wealthy employees are relatively more sophisticated than unwealthy ones in adjusting to market exposures when making investment decisions. Utkus and Young (2011) find participants with low financial test scores borrow more from their 401(k) plans. Benartzi and Thaler (2001, 2007) find financial illiteracy is linked with poor diversification decisions. We then construct three measures based on their 401(k) loan activity and asset diversification and show that our wealthy employees are indeed more financially sophisticated, as they borrow less, pay off loans faster, and have more diversified investments than unwealthy ones. Moreover, the return-chasing behavior of the unwealthy group exists in both active funds and index funds, while the wealthy group does not have any flow-performance sensitivity among index funds, further supporting that

[2] In the 401(k) context, deferral rate is the portion of wages automatically deducted from employee's paycheck and invested in the 401(k) account each pay cycle.

the wealthy group is more financially sophisticated.

Lastly, we show that the unwealthy employees walk away from substantial capital gains by chasing returns. We consider a simple passive strategy, in which participants allocate flows to equity and bond funds following a pre-determined cutoff, for example, 80% in equity and 20% in bond. Such a strategy is similar to investing in target date funds, except that the asset allocation gradually changes over time in target date funds. When there are multiple equity (bond) funds in the plan, flows are evenly invested in those funds, which effectively shuts down any fund selection from participants' perspective. We show that, for a wide range of asset allocation ratio, our passive strategies significantly outperform the observed investments of participants by a large margin. Our result suggests that unsophisticated investors are better off avoiding active strategies and investing passively.

Our paper contributes to the literature that studies asset allocation decisions of households in their retirement accounts. For example, Madrian and Shea (2001) and Agnew, Balduzzi, and Sunden (2003) find there is inertia in asset allocation. Calvet, Campbell, and Sodini (2007) study household's diversification using Swedish data and find that they incur return losses from under-diversification. Benartzi and Thaler (2001) find investors follow a naive "1/n" strategy in their retirement accounts. Unlike prior studies, our paper investigates which performance measure investors respond to in their retirement accounts by using a large and comprehensive dataset covering all public firms' 401(k) plans between 1993 and 2016. Sialm et al. (2015) and Pool et al. (2016) focus on the role of plan sponsors in driving flows. Our paper shows that participants and plan sponsors react to different performance signals.

Our paper also contributes to the large literature that has examined flow-performance relations almost exclusively at the aggregated fund level, which is mainly driven by large investors. Barber et al. (2016) and Berk and van Binsbergen (2016) find flows respond to CAPM alpha, while Evans and Sun (2018) and Ben-David et al. (2019) find flows respond to Morningstar rating. Rather than focusing on the investment behavior of large investors in the economy, our paper tries to understand the investment pattern for individual investors in their 401(k) plans. Even though we observe only the overall holding at the plan level, the effect of wealthy individuals is considerably weakened because of the 401(k) contribution cap and the large number of employees in a firm. Therefore, our analyses speak for the pattern of the majority of individual investors in their retirement accounts. Different from prior findings, we show that the majority of individual investors simply follow fund returns.

The rest of this study is organized as follows. Section 2 describes the dataset and constructions of key performance measures. Section 3 shows the flow-performance relation at the plan level. Section 4 discusses the mechanisms that drive flow-performance relation. Section 5 concludes our discussion.

# 2 Data

## 2.1 Data sources

We hand-collected employees' investments in 401(k) plans from the annual Form 11-K, which is required by the U.S. Securities and Exchange Commission (SEC) if a firm offers company stock to employees in its plan. Our sample includes 1,551 public firms from 1993 to 2016. Firms report a list of securities available and the current investment value for each security. The list typically consists of the employer's stock and a set of mutual funds. For a few firms that provided multiple 401(k) plans for different subsidiaries in a given year, we aggregated them into one plan. We supplement our dataset by collecting Form 5500 from the Department of Labor. Firms are required to report plan-level information, such as the amount of employee and employer contributions, assets value, loans to participants, loan interest, and whether the plan is defined contribution or defined benefit. Around 70% of our 11-K data can be merged with Form 5500 data using the firms' Employer Identification Number (EIN).

The fund-level data are from the Center for Research in Security Prices (CRSP) survivorship bias-free mutual fund database and Morningstar database. We matched funds in CRSP with the ones in the 11-K by fund name. Since most funds are listed on 11-K without share class information, we aggregated share-class-level data in CRSP to the fund level. For example, fund expense ratios and returns at the share-class level are aggregated using their previous month's total net assets (TNA) as weights. Fund TNA are the sum of TNA of all share classes within the same fund.

The firm-level data are from CRSP and Compustat. For example, we obtain the Standard Industrial Classification (SIC) code, headquarter locations, and the number of employees of a firm from Compustat.

## 2.2 Sample statistics

Our sample covers a wide range of firms with a balanced distribution across industries. The distribution of firms in our sample across the Fama-French 12 industries is very close to that of all public firms. Within each industry, our sample covers 12% of firms on average. This number varies from 9% to 15% for all industries except for the utility (32%), healthcare (5%), and business equipment (6%) industries. Moreover, our sample tends to cover large firms. For example, 80% of our firms are larger than the industry median, and 40% are in the top size quintile.

Table 1 provides descriptive statistics by year at the plan level. The average and median plan sizes are \$380 million and \$68 million, respectively, and they increase over the years except during the recessions of 2001 and 2007-2009, which caused depreciation in plan value. In our sample, the total investment in 401(k) accounts has grown from \$37 billion in 1993 to around \$300 billion in 2016. The increasing number of firms that offer 401(k) plans contributes to this trend. Moreover, our sample covers an average of 8% of plan assets across all public and private firms that have plans with 100 or more participants.[3]

Employees invest substantially in their firm's stock in the early years of our sample.[4] The shares invested in company stock have declined gradually from 48% in 1993 to 24% in 2016, along with an increase in the number of available funds. For example, an average of 17 funds are offered in 2015 compared to only three funds in 1993.[5] We then split funds into three categories: equity funds, bond funds, and a blend of these two fund types. Within funds, employees allocate 67% of their capital to equity funds and 20% to bond funds. These allocations are stable over the years examined.

## 2.3 Performance measurement

In this section, we discuss four performance measures used in this paper: [1] fund's net-of-fee return ($R_{ft}$), [2] CAPM alpha ($\alpha_{ft}^{CAPM}$), [3] 4-factor alpha ($\alpha_{ft}^{4Factor}$), and [4] Morningstar return ($MStar\ return_{ft}$).

The CAPM and 4-factor alphas are estimated monthly based on a rolling estimation window following Barber et al. (2016). For each fund $f$ in month $m$, we estimate the

[3] Data on 401(k) assets for all firms are in Private Pension Plan Bulletin from the Department of Labor: `https://www.dol.gov/agencies/ebsa/researchers/statistics/retirement-bulletins/private-pension-plan`.

[4] The under-diversification is also documented in Mitchell and Utkus (2002), and Poterba (2003).

[5] This number is smaller in 2016 because fewer firms are included in the dataset for this year due to firms likely filing late.

following time-series regressions using 36 months of returns:

$$R_{f\tau} - RF_\tau = a_{fm} + \mathbf{F}'_\tau \boldsymbol{\beta}_{fm} + \varepsilon_{f\tau}, \quad \tau = m-1, ..., m-36 \tag{1}$$

where $RF_\tau$ is the risk-free rate in month $\tau$, and $\mathbf{F}_\tau$ is the vector of factor returns. For equity and balanced funds, we use the CRSP value-weighted stock index (equity market) factor in the CAPM model and use Carhart (1997) 4-factor model that includes market, size, value, and momentum factors. For bond funds, we use the U.S. aggregate bond index (bond market) in the CAPM model, and use the following four factors in the 4-factor model: the equity market index, the bond market index, the U.S. high-yield bond index, and the mortgage-backed security index. These factors have been used in Ma, Tang, and Gomez (2019), Cici and Gibson (2012), and Elton, Gruber, and Blake (1995). We then estimate alpha for fund $f$ in month $m$ as follows:

$$\hat{a}_{fm} = R_{fm} - RF_m - \mathbf{F}'_{\mathbf{m}} \hat{\boldsymbol{\beta}}_{fm}, \tag{2}$$

where $\hat{\boldsymbol{\beta}}_{fm}$ is estimated from equation (1). Since our data are at an annual frequency, an annual alpha of fund $f$ is calculated using the 12 monthly alpha in year $t$:

$$\alpha_{ft} = \prod_{j=0}^{11} \left(1 + \hat{a}_{f,t-\frac{j}{12}}\right) - 1. \tag{3}$$

The fund's net-of-fee return ($R_{ft}$) is annualized in the same way.

To assign a star rating, Morningstar uses the risk-adjusted return $MRAR$ to rank fund[6]

[6]The Morningstar Rating for Funds is available at https://s21.q4cdn.com/198919461/files/doc_downloads/othe_disclosure_materials/MorningstarRatingforFunds.pdf. Before 2002, Morningstar uses a different formula for risk-adjusted return, which is described in Blume (1998), to rank funds. Therefore, we use accordingly, and only a small fraction of our sample (10%) falls before 2002.

$f$ at time $t$ as follows:

$$MRAR_{ft}(\gamma, T) = \left[\frac{1}{T}\sum_{j=0}^{T-1}(1 + ER_{f,t-j})^{-\gamma}\right]^{-\frac{12}{\gamma}} - 1, \tag{4}$$

where $\gamma$ is the risk aversion coefficient and $ER_{f,t-j} = \dfrac{R_{f,t-j} - RF_{t-j}}{1 + RF_{t-j}}$ is the geometric return in excess of the risk-free rate. Morningstar uses $\gamma = 2$ to rank funds. We use the most recent 36 months of return to compute $MRAR_{ft}(2, 36)$, which is the Morningstar annualized return for the fund at year $t$. Next, the Morningstar return ($MStar\ return_{ft}$) used throughout this paper is the standardized value of $MRAR_{ft}(2, 36)$ within each investment category, and it reflects the fact that Morningstar assigns a star rating to each fund relative to other funds in a given investment category (i.e., the top 10% will be 5-star funds, and the bottom 10% will be 1-star funds).

# 3 Return chasing in 401(k) plans

Even though both employees and employers contribute to 401(k) retirement plans, employees are responsible for making their own investment decision. Therefore, our dataset provides a perfect context in which to study how investors evaluate fund performance and make investment decisions accordingly in their retirement accounts. Moreover, our dataset tracks the menu of funds in a 401(k) plan, so that we can further investigate the flow-performance sensitivity when the menu changes.

We use a variety of methods to examine what performance measures investors use to direct their flow of capital, which is defined as follows:

$$Flow_{pft} = \frac{V_{pft} - V_{pf,t-1}(1 + R_{ft})}{V_{p,t-1}}, \tag{5}$$

where $V_{p,t-1} = \sum_{f \in \Theta_{p,t-1}} V_{pf,t-1}$, and $V_{pft}$ is the investment value in fund $f$ from the participant of firm $p$'s 401(k) plan in year $t$ and $R_{ft}$ is the fund's net-of-fee return during year $t$. $\Theta_{p,t-1}$ is the set of funds in firm $p$'s plan in year $t-1$, hence the denominator represents the firm's plan size in that year. We scale flows by the plan value, as we are interested in investors' within-plan allocation.[7] To remove outliers driven by potential data entry errors, we winsorize the variable *Flow* at 1% and 99% levels.

## 3.1 Panel regression

We first study the flow-performance sensitivity by estimating the following regression:

$$Flow_{pf,t+1} = \beta_0 PERF_{ft} + \mathbf{X'}_{ft}\boldsymbol{\beta}_1 + \mu_{pt} + \epsilon_{pf,t+1}, \tag{6}$$

where $PERF_{ft}$ is fund $f$'s performance measures in year $t$. There are four performance measures used in this paper: [1] fund's net-of-fee return, [2] CAPM alpha, [3] 4-factor alpha, and [4] Morningstar return. Detailed construction is discussed in Section 2.3. To focus on investors' flow-sensitivity relation within their 401(k) plan in a given year, we add firm×year fixed effect ($\mu_{pt}$). The fixed effect absorbs not only variations across firms, such as the industry that the firm is in, but also time-varying shocks to the firm, such as product-market shocks, which may have an effect on employees' investment decisions. We also control for a set of fund characteristics, $\mathbf{X}_{ft}$, including fund expense ratio, fund turnover ratio, the

[7]The measure is the same as the flow measure *NMG3* in Pool et al. (2016).

logarithm of total fund net assets, standard deviation of fund returns in the past year, and investment style fixed effect. One important feature of our regression is that we study within-plan flow-performance relations and shed light on the majority of investors' asset allocation decisions. Our setting contrasts with the typical fund-level flow-performance regression that studies the representative agent's decisions.

Table 2 shows the regression estimation results. Each performance measure can predict future within-plan flows on its own, but net return clearly wins all possible horse-race with other performance measures. In columns (1) to (4), each of these four performance measures in year $t$ significantly predicts fund flows in year $t+1$. In Appendix Table A1, we also add performance measures before year $t$, and flows only respond to the most recent performance.

To compare predictive powers across models, we first make pairwise comparisons between the fund's net return and other performance measures. Column (5) reports the regression results of future flows on net return and CAPM alpha. The coefficient on the CAPM alpha not only becomes much smaller, but it is also insignificant, while the coefficient of net return is 0.018 and is statistically significant at the 5% level. The results are similar when comparing net return with the 4-factor alpha or Morningstar return in columns (6) and (7). Lastly, we run the horse-race among these four performance measures. Results in column (8) show that net return is the winner while the others do not drive future flows. The economic magnitude of the net return coefficient is also large. The median fund size relative to a 401(k) plan is around 5%. Therefore, a 1% increase in a fund's net return corresponds to a 0.02% increase in flows relative to the plan size, which is equivalent to $0.02\%/5\% = 0.4\%$ extra flow to the fund.

## 3.2 Robustness check: nonparametric approach

In the previous section, we examined flow-performance sensitivity assuming a linear relation. To relax the linear relation assumption, we employ two additional robustness checks, following Barber et al. (2016) and Ben-David et al. (2019). Both tests show that, consistent with the linear regression result, net return is the winning performance measure that the majority of investors use in our sample.

### 3.2.1 Pairwise model horse race

Similar to Barber et al. (2016), we take advantage of the heterogeneous rankings of a fund within a firm due to different performance measures by conducting a pairwise horse-race between net returns and other measures. Specifically, we run the following regression:

$$Flow_{pf,t+1} = \sum_i \sum_j b_{ij} D_{ijpft} + \mathbf{X'}_{ft}\mathbf{c} + \mu_{pt} + \zeta_{pf,t+1}, \tag{7}$$

where $D_{ijpft}$ is a dummy variable that equals one if fund $f$ of firm $p$ in year $t$ is in tercile $i$ based on the net return measure and tercile $j$ based on an alternative performance measure, such as the CAPM alpha or the Morningstar standardized return. Similar to the equation (6), we also include firm×year fixed effect $\mu_{pt}$ and a set of fund-level controls $\mathbf{X}_{ft}$. Note that plans with fewer than three funds are dropped, which accounts for 1% of the observations.[8]

To estimate the model, we exclude the dummy variable for $i = 2$ and $j = 2$. By excluding the dummy variable, $b_{ij}$ represents the percentage flows to a fund in tercile $i$ and $j$ based on fund net return and other return measures, respectively, relative to a fund that was in the middle tercile for both performance measures. Investors use net returns to direct their flows

[8] Our results are robust and stronger when we rank funds into quintiles. Under the quintile specification, plans with fewer than five funds are dropped, which accounts for 5.5% of the observations.

if the sum of differences in coefficients on dummy variables $(b_{i,j} - b_{j,i})$ for all $i$ and $j$ such that $i > j$:

$$\text{Sum} = \sum_{i>j}^{3} \sum_{j=1}^{2} (b_{i,j} - b_{j,i}) \tag{8}$$

is significantly greater than zero.

Table 3 reports the sum of differences in coefficients. The results show that, consistent with the result from linear regression, flows respond to funds with high returns rather than CAPM alphas, 4-factor alphas, or Morningstar standardized returns, as the sums of coefficients are all positive and significant.

#### 3.2.2 Top-ranked and bottom-ranked funds

Following Ben-David et al. (2019), we test which performance measure has the greatest power in explaining within-plan variation in flows, both in terms of magnitude and direction. Within each plan, we rank funds into terciles by returns, CAPM alpha, 4-factor alpha, and Morningstar standardized returns,[9] respectively. We then compare flow differences between the top and bottom-tercile funds. Specifically, we examine whether the difference in performance ranking can explain the fraction of funds with positive flows, flows in percentage points, and flows in dollars.

The result is shown in Table 4. Consistent with previous results, net return has the most explanatory power for fund flows in positive flows, percentage points, and dollars. For example, the spread in return ranking generates a percent flow gap of 0.71%, which is 40% greater than what the spread in CAPM alpha or Morningstar rating generates. Moreover, the spread in return ranking generates a dollar flow gap of $565k, which is 72% (290%)

[9] Our results are robust if we use Morningstar's star ratings instead of Morningstar standardized returns to rank funds.

greater than what the spread in CAPM alpha (Morningstar rating) generates.

# 4 Mechanism of return chasing

Having documented the return chasing in 401(k) plans, we answer why our results differ from previous literature and explore mechanisms that drive the return chasing behavior in this section.

First, previous literature finds that the representative agent in the economy uses CAPM alpha or Morningstar ratings to make investment decisions, where the observations are at the fund level. In contrast, our analyses in the previous section focus on flow-performance sensitivity at a more granular level than the previous literature, and we aim to provide evidence on whether the majority of firm employees focus on the fund returns. Even though the scope of questions is different between our paper and the previous literature, one alternative explanation for the differing results could be that investors in the CRSP sample are very different from ours. To rule out this explanation, we aggregate our plan-level data to the fund level and re-examine flow-performance sensitivity in section 4.1.

Second, the investment decision in 401(k) plans can be viewed as a two-step process. In the first step, employers and plan sponsors create the menu. Whether to add or delete a fund depends on its recent performance and its affiliation with the sponsor (Pool et al. (2016)). In the second step, employees observe the menu and make their investment decisions. At the aggregate level, if a fund is added by multiple 401(k) plans simultaneously, the increased exposure will attract more flows, holding all else equal. Mechanically, there will be a positive flow-performance relation driven by the menu effect, as the added fund tends to perform well in the past. Since our paper focuses on flow-performance relation from investors' perspective, we decompose fund flows into sponsor flows and participant flows, following Sialm et al. (2015), to address the effect of the menu change. Moreover, we also test the robustness of our plan-level findings at two subsamples, one with and one without menu change. The results are detailed in section 4.1.

Third, we explore mechanisms that drive the difference between fund-level and plan-level results. We study whether relatively wealthier employees follow a more sophisticated performance measure in section 4.2. The level of wealth not only has a heterogeneous effect on flow-performance sensitivity in mutual funds investment, but also varies with how employees make investment decisions between funds and company stock. We are the first to directly test how the flow composition between company stock and funds changes with respect to their performance gap. The results are shown in section 4.3.

Fourth, we link the variation in 401(k) savings to financial literacy. We show that common determinants of 401(k) savings, such as employee age, firm's defined contribution matching policy, and plan performance, can only explain a small fraction of variations in savings, leaving a large fraction to be explained by employees' wage and participation rate, which highly correlates with financial literacy. In section 4.4, by constructing three measures of financial literacy based on employees' loan activity and asset allocation at the firm-year level, we show that the wealthy group indeed has higher measures of financial literacy than the unwealthy group. Consistent with this finding, we show that the unwealthy group cannot distinguish between active and index funds and exhibit return-chasing behavior in both types of funds, whereas the wealthy group does not show any flow-performance relation among index funds.

Lastly, in section 4.5, we show that unwealthy employees incur a significant loss when chasing net-of-fee returns, as their investments largely underperform a passive strategy that invests in the existing funds in their 401(k) plans.

## 4.1 Menu change and flow-performance sensitivity

To reconcile the difference between our plan-level results and the results in the literature (Barber et al. (2016), Berk and van Binsbergen (2016)), we aggregate plan-level data to the fund level and re-examine the flow-performance relation. Specifically, the total flow of new money to fund $f$ from all 401(k) plans in year $t$ is

$$\textit{Total Flow}_{ft} = \frac{V_{ft} - V_{f,t-1}(1 + R_{ft})}{V_{f,t-1}}, \tag{9}$$

where $V_{ft} = \sum_{p \in \Omega_{ft}} V_{pft}$ is the total investment value to fund $f$ from employees of all firms in year $t$. We then run the following panel regression of aggregated flow of fund $f$ in year $t+1$ on performance measures in year $t$, controlling for fund characteristics and time fixed effects as follows:

$$\textit{Total Flow}_{f,t+1} = \delta_0 PERF_{ft} + \mathbf{X'}_{ft}\boldsymbol{\delta}_1 + \kappa_t + \eta_{f,t+1}, \tag{10}$$

where $\mathbf{X'}_{ft}$ represents fund characteristics which are a vector of control variables used in the equation (6). $\kappa_t$ is year fixed effect.

The results are shown in column (1) of Table 5. At the aggregate level, both CAPM alpha and Morningstar return can predict fund flows, which is consistent with the prior literature. However, it is too early to conclude that investors direct flows according to the CAPM model or Morningstar ratings. For example, the menu change can impact the aggregated fund flows (Sialm et al. (2015)). Specifically, plan sponsors may add funds that perform well in the past or have high Morningstar ratings and delete funds with bad performance or low ratings

(Pool et al. (2016)), which can mechanically produce a flow-performance sensitivity.

To tease out the effect of the sponsor's menu change on fund flows, we then decompose the aggregated fund flows into the *Sponsor Flow* and *Participant Flow*, following Sialm et al. (2015). Specifically, we define:

$$Sponsor\ Flow_{ft} = \frac{\sum\limits_{p \in Additions} V_{pft} - \sum\limits_{p \in Deletions} V_{pf,t-1}(1 + R_{ft})}{V_{f,t-1}}, \tag{11}$$

$$Participant\ Flow_{ft} = Total\ Flow_{ft} - Sponsor\ Flow_{ft}. \tag{12}$$

We then regress *Sponsor Flow* and *Participant Flow* in year $t+1$ on performance measures in year $t$, controlling for fund characteristics and time fixed effects. Columns (2) and (3) of Table 5 show the regression results. Both the sponsor flow and participant flow follow the CAPM alpha, which rules out the alternative explanation that the cross-sectional variation in flows is only driven by the change in menus.

The menu change may also affect flows at the plan level. For example, the firm and fund advisors may use a different performance metric to select or delete funds, which will lead to a reduction in the sensitivity of our observed flow-return relation. To examine the menu effect, we decompose our plan-level flows into three parts: flows to newly added or deleted funds, which we also denote as the sponsor flow; flows to remaining funds when there is a menu change; and flows to remaining funds when there is no menu change. As shown in column (4) of Table 5, sponsor flows are positively correlated with Morningstar returns, consistent with the aggregated-level evidence. Meanwhile, participant flows to remaining funds only respond to net returns regardless of whether the menu has changed, which can be seen from the positive coefficient on net returns in columns (5) and (6) of Table 5.

## 4.2 Reconciling the aggregated and plan-level analyses

Why do aggregated participant flows respond to the CAPM alpha, while the plan-level flows respond to net returns? We conjecture that the difference is driven by different weighting schemes. In particular, the plan-level evidence suggests that the "average" employee in the firm allocates savings by following returns. The analysis effectively puts more weight on firms with small plan sizes. In contrast, the aggregated analysis puts more weight on firms with large plan sizes. We offer a potential explanation through the wealth channel to reconcile the differences in results between our research and prior studies. In particular, we hypothesize that a small proportion of employees possess a substantial portion of savings, and their investment decision follows the CAPM alpha.

To test this hypothesis, each year, we rank firms into quartiles by the average savings per employee of the plan. We use average savings instead of flows because the latter is more subject to retirement withdrawal and more volatile. Moreover, the average savings in a given 401(k) plan are less likely to be affected by a small set of wealthy individuals, since there is a contribution limit, i.e. the participant's maximum deferral amount was $10,500 in 1994 and $18,000 in 2016, and our sample focuses on large public companies with over 17,000 employees on average. For each quartile, we re-run the plan-level panel regression of flows on performance measures. The results are shown in Table 6. Almost all quartiles show a flow-return relation, except for the highest-savings quartile, where flows respond to the CAPM alpha. Note that we do not argue that, in practice, wealthy employees pull out historical returns and estimate the CAPM alpha. One possible explanation could be that they focus on benchmark-adjusted returns, which will be highly correlated with CAPM alpha, given that many funds have a market beta close to one and use S&P 500 index as their benchmark.

Table 7 examines the plan characteristics of these subsamples. Employees in the highest-savings group (hereafter, *wealthy group*) save on average \$121,616 per person, whereas the remaining employees (hereafter, *unwealthy group*) only save \$15,592. The wealthy group also has significantly fewer employees than the unwealthy group. The average plan size of the wealthy group is \$840 million, while that of the unwealthy group is \$201 million.

Figure 1 visualizes the gap in savings between the wealthy and unwealthy groups. In each year, we rank firms by the average savings per employee into percentiles, and calculate the cumulative fractions of the number of employees and total 401(k) savings, respectively. For each percentile, we then average the two cumulative fractions over the years. Similar to our regression analyses, wealthy participants are defined as the ones ranked in the top quartile (75th percentile), the cutoff of which is denoted by the red vertical line. In fact, the wealthy group represents only 17% of the employees in our sample, but holds more than 61% of the total savings. As a result, the flow-performance relation is mainly driven by the wealthy group at aggregated fund level. Meanwhile, the majority's flow-performance sensitivity stands out when we estimate the relation at the more granular plan level.

There are two potential concerns about whether our average savings capture the wealth level. The first concern is that some firms may also offer defined benefit plans. If our classified unwealthy group is more likely to have defined benefit plans than the wealthy group, we may underestimate the wealth level for the unwealthy group and overestimate the wealth level for the wealthy group. The second concern is that some employees may keep 401(k) plans with the previous employer, and our test implicitly assumes employees always transfer their savings into the new plan when they move. In Appendix A, we discuss these two concerns in detail and show that our results are not affected.

### 4.3 Flow gap between company stock and mutual funds

The return-chasing behavior not only exists when employees make investment decisions within mutual funds but also between the company stock and mutual funds. In this section, we focus on the measure of *Flow gap*, which is the difference between employees' dollar flows to their company's stock ($\$Flow_{pt}^{stock}$) and mutual funds in the 401(k) plans ($\$Flow_{pt}^{funds}$), normalized by the total investment value to company stock and all funds in plan $p$ ($V_{p,t-1}^{s}$):

$$Flow\ gap_{pt} = \frac{\$Flow_{pt}^{stock} - \$Flow_{pt}^{funds}}{V_{p,t-1}^{s}}. \tag{13}$$

The mean, standard deviation, and kurtosis of *Flow gap* are -0.06, 0.19, and 8.36, respectively. The negative mean suggests that, unconditionally, there are more flows to mutual funds than to company stock, which is consistent with mutual funds representing roughly 70% of 401(k) assets. The large standard deviation and kurtosis imply the large cross-sectional variations.

We then study whether employees' asset selection in 401(k) depends on the difference in past performance between company stock and funds, which we denote as $\Delta PERF_{pt}$.

$$\Delta PERF_{pt} = PERF_{pt}^{stock} - PERF_{pt}^{funds}, \tag{14}$$

where $PERF_{pt}^{stock}$ is the performance of the company stock, and $PERF_{pt}^{funds}$ is the performance of mutual funds in the plan, weighted by the investment value. We measure performance in three ways: [1] net return, [2] CAPM alpha, and [3] 4-factor alpha. We then estimate the following regression:

$$Flow\ gap_{p,t+1} = \theta_0 \mathbb{1}\{\Delta PERF_{pt} > 0\} + \mathbf{Z'}_{pt}\boldsymbol{\theta}_1 + \lambda_p + \kappa_t + \iota_{p,t+1}, \tag{15}$$

where $\mathbb{1}\{\}$ is a dummy variable which equals one if the performance gap is positive and zero otherwise; $\mathbf{Z}'_{pt}$ represents a set of firm and fund characteristics, including the natural logarithm of firm size, book-to-market, the logarithm of the number of funds in the menu, the logarithm of company stock and fund size in 401(k) plan, funds' weighted expense ratio, turnover ratio, and the natural logarithm of the sum of fund size in the plan. $\lambda_p$ and $\kappa_t$ are firm and year fixed effects, respectively.

Table 8 reports the regression result. Consistent with the return-chasing behavior within mutual funds, the unwealthy group allocates more flows to the company stock when the company stock has a higher return than mutual funds in the plan, whereas the wealthy group follows the difference in CAPM alpha.

## 4.4 Financial literacy and 401(k) savings

As we have shown in the previous sections, employees of the wealthy group save an average of $121,616, which is over 7.8 times greater than the average savings of the unwealthy group. What factors contribute to the difference in savings? There are six main determinants: the return of plan assets, firm's matching contribution ratio, firm's matching contribution cap, employees' working tenure, employees' salary, and employees' deferral rate. In Appendix B, we show that 75% of the savings gap is left to be explained by employees' salary and deferral rate. We argue in this section that financial literacy is strongly related to these two factors and plays an important role in explaining the savings gap. Moreover, financial literacy can also help to explain why wealthy employees in our sample are relatively more sophisticated than unwealthy ones when making investment decisions, as their money flows mainly respond to the CAPM alphas, hence taking into account market exposure.

Figure A1 provides some suggestive evidence on the link between financial literacy and the level of savings. We use the State-by-State survey data from the Foundation's National Financial Capability Study (FINRA). These surveys have been conducted since 2009 to study the financial capability of American adults. The dataset contains not only respondent characteristics, such as ages, education levels, and household income, but also various variables pertaining to how American adults make financial decisions. The surveys also include a financial literacy test that asks respondents five finance questions. As shown in the figure, respondents with higher test scores are more likely to make regular contributions to their retirement plans and have higher incomes.

To formally test whether our wealthy employees indeed are more financially sophisticated than the unwealthy ones, we construct three measures to proxy for financial literacy at the firm-year level using Form 11-K and Form 5500 data.

The first two measures use the information on loans to 401(k) participants and the associated loan interest. Utkus and Young (2011) conduct a survey of individual plan participants from Vanguard's 401(k) record-keeping system and find that participants who scored low on questions relating to general financial knowledge are more likely to borrow from their 401(k). Therefore, our first financial (il)literacy measure is the loans to the participants scaled by the 401(k) plan size. The second measure is the loan interest scaled by the amount of loans. Conditional on the same amount of loans, the loan interest can capture how early participants tend to pay off their loans. The third measure is the flow to company stock. Benartzi and Thaler (2007) argue that investing in the employer's stock is a sign of poor diversification, as a single security is much riskier than a portfolio of mutual funds, and the labor income is positively correlated with the stock performance.

Note that our measures should be negatively correlated with financial literacy. To test whether wealthy employees have higher financial literacy, we then regress the wealthy dummy on these three measures, controlling for the main determinants of employee savings in 401(k) (firms matching contribution ratio, firms matching contribution cap, employees working tenure, return of plan assets, and employee contribution) and firm characteristics. Specifically, we estimate the following regression:

$$\begin{aligned} Wealthy_{pt} = {} & \beta_1 \mathit{Finance\ literacy}_{pt} + \beta_2 \mathit{Firm\ matching\ ratio}_{pt} + \beta_3 \mathit{Firm\ contribution\ cap}_{pt} \\ & + \beta_4 \mathit{Firm\ age}_{pt} + \beta_5 \mathit{Plan\ cumulative\ return}_{pt} + \beta_6 \mathit{Employee\ contribution}_{pt} \\ & + \mathbf{M'}_{pt} \boldsymbol{\beta}_7 + \kappa_t + \xi_{pt}, \end{aligned} \tag{16}$$

where $Wealthy_{pt}$ is a dummy indicator that equals one if the firm is ranked in the top quartile by the average savings. $\mathit{Firm\ matching\ ratio}_{pt}$ and $\mathit{Firm\ matching\ cap}_{pt}$ are the ratio of the firm's contribution to the employee's contribution and the maximum amount that a firm is willing to contribute as a percentage of an employee's annual compensation. We hand-collected this data from form 11-K. For missing observations of $\mathit{Firm\ matching\ ratio}_{pt}$, we use data from Form 5500. $\mathit{Firm\ age}_{pt}$ is the logarithm of firm age that proxies for employees working tenure. $\mathit{Plan\ cumulative\ return}_{pt}$ is value-weighted cumulative returns of all securities in 401(k) plan. $\mathit{Employee\ contribution}_{pt}$ is the logarithm of an annual 401(k) contribution per employee. The data on employee contribution is from Form 5500. Note that employee contribution is the product of employee salary and deferral rate, and it captures the joint effect on wealth accumulation. Summary statistics of these variables in the wealthy and unwealthy groups are reported in Table 7. $\mathbf{M'}_{pt}$ represents a set of firm characteristics, which includes the logarithm of market capitalization, book-to-market ratio, debt-to-asset ratio, gross profitability ratio, and a dummy of whether a firm has a defined benefit plan. $\kappa_t$ is the year fixed effects.

The regression result is shown in columns (1) to (3) of Table 9. After controlling for common factors that determine the savings, we find that employees in the wealthy group have higher financial literacy than the unwealthy group. The wealthy group tends to borrow less from 401(k) plans, pay off their loans earlier, and invest less in company stock than the unwealthy group.

Next, we show that our identified gap in financial literacy between wealthy and unwealthy groups is not concentrated in certain industries. First, we tabulate the relative frequency of each Fama-French-12 industry in our wealthy and unwealthy groups separately. The results in the first two columns of Table 10 show that the industry distributions do not differ much between the two groups, except for a few industries. Energy and utility industries are more likely to appear in the wealthy group, while manufacturing and retail industries are more likely to appear in the unwealthy group.

Second, we test the difference in our three measures of financial literacy between wealthy and unwealthy groups by industry. The last three columns of Table 10 show the difference and statistical significance, and all numbers are in percentage points. When financial literacy is measured by the fraction of loans and flows to company stock, our wealthy classification can capture variations in financial literacy for all industries, as all the numbers in these two columns have the expected sign, and almost all are significantly different from zero. When financial literacy is measured by loan interest, only the health industry has a contradictory positive and significant difference at 10% level. Our results suggest that, although variation across industries can capture some differences in financial literacy, there are also within-industry variations in financial literacy that can be captured by our wealthy classification.

Third, we re-estimate the flow-performance regression in the wealthy group by omitting industries one-by-one and present the results in Table A2. The coefficient estimates on the CAPM alpha are very stable, regardless of the omitted industries, suggesting that our finding is not concentrated in any particular industries.

To formally control for the industry-level variations in wealth accumulation, financial literacy, and compensation, we re-estimate Equation (16) by including the industry $\times$ year fixed effects. This allows us to focus on within-industry variations in wealth accumulation and how it relates to the financial literacy. The regression results are shown in columns (4) to (6) of Table 9. Consistent with the baseline results, financial literacy plays a role in determining the savings gap between the wealthy and unwealthy groups, as the coefficient estimates of all three financial illteracy measures are significantly negative.

We also conduct a set of robustness checks. First, we show that our results are robust in subsamples of firms with and without defined benefit plans (Table A3). Second, the savings per employee also depends on employee locations and how easy it is to switch jobs. For example, the cost of switching jobs will be much lower for a software engineer located on the west coast than in the central region. To account for unobservables at the industry and geographical location level, we use industry $\times$ state fixed effects in Table A4, and the results are very similar to the baseline regressions. Third, one may concern that firm age may not capture the time of employees' wealth accumulation, so that we use the inferred employee age based on the year of Target Date Funds in the plan as a proxy. The results remain robust (Table A5).[10]

Having documented that the wealthy group is more financially sophisticated than the unwealthy group, we then test whether their flow-performance relation differs between active and index funds. If the unwealthy group is indeed financially unsophisticated, we should expect there is also return-chasing behavior among index funds. On the other hand, the wealthy group should not have flow-performance sensitivity for index funds.

[10]Appendix D details on how employee age is inferred.

To test this channel, we regress fund flows on various performance measures, an index fund dummy, and their interaction terms, controlling for firm×year fixed effects and fund characteristics. Taking CAPM alpha as an example, the coefficient estimate of CAPM alpha captures the sensitivity of active funds, while the sum of the coefficient estimates of CAPM alpha and its interaction term with the index fund dummy captures the sensitivity of index funds. The regressions are estimated separately for wealthy and unwealthy groups. Table 11 presents the flow-performance sensitivity for active funds and index funds. As shown in the first two columns, employees in the unwealthy group have positive flow-return sensitivities for both index and active funds, consistent with them being less financially sophisticated. Employees in the wealthy group, however, do not have any flow-performance sensitivity for index funds, and invest in active funds based on CAPM alpha.

## 4.5 Economic loss of return-chasing

Lastly, we show that employees in our unwealthy group give up substantial capital gains by chasing net-of-fee returns, and they are better off by allocating savings passively to funds in their plans.

For simplicity, we consider a passive allocation strategy with an allocation rule that invests, for example, 80% in equity and 20% in bonds.[11] At the end of each year and for each plan, given the total dollar flows, we redistribute 80% of the flows to the equity funds and 20% to the bond funds in the plan. If the plan has more than one equity (bond) funds, we distribute flows equally among them. The results are very similar if we use value weighting rather than equal weighting. We then examine the performance difference between the observed flows and the hypothetical flows using such passive strategy in Figure 2. We

[11] For simplicity, we group balanced funds into bond funds. Results do not change if we group balanced funds into equity funds.

use three asset allocation cutoffs, 80%, 60%, and 40% in equity, denoted by the green, red, and black lines, respectively. Note that this passive strategy is similar to investing in a target-date fund, except that the asset allocation becomes more conservative over time in a target-date fund. For example, target-date funds with a target retirement in 20, 10, and fewer than five years will invest 80%, 60%, and 40% in equity, respectively. However, as we show in the figure, even if we consider an allocation that invests 40% in equity, which is very conservative, it still outperforms the observed flows of unwealthy investors by a large margin.

Panel (a) of Figure 2 studies the short-term performance difference between the two types of flows. We calculate the returns of flows at year $t$ in year $t+1$, and then plot the cumulative returns over the years in our sample period. The passive strategies with 80%, 60%, and 40% equity allocation have cumulative returns (Sharpe ratios) of 512% (0.56), 456% (0.52), and 400% (0.49) over the sample period. The observed flows significantly underperform the passive strategies. Specifically, the observed flows earn around 237% with a Sharpe ratio of 0.30. The difference in dollar terms is about $1 billion over the sample period for the passive strategies with 60% equity allocation.

Therefore, the unwealthy employees are worse off chasing funds with high returns in the past. The intuition is that the past performance of mutual funds is a poor indicator of managers' skill (Barras, Scaillet, and Wermers (2010)). When picking funds based on past returns, funds that are bought can potentially underperform funds that are sold by the unwealthy investors, whereas the passive strategy allows investors to buy these funds, regardless of their recent performance. To test this channel, we decompose the performance difference between the observed flows and hypothetical flows into two components. The first component, denoted by $R_{t+1}^{opposite}$, comes from the scenario in which observed flows and hypothetical flows have opposite directions, which are often a result of investors selling recent underperforming funds. The second component, denoted by $R_{t+1}^{same}$, arises when the two flows have the same signs and only differ in magnitude. Specifically,

$$R_{t+1}^{Passive} - R_{t+1} = R_{t+1}^{same} + R_{t+1}^{opposite} \tag{17}$$

$$R_{t+1}^{same} = \sum_{f \in \Omega_t} \sum_{p \in \Omega_{ft}} (w_{pft}^{Passive} - w_{pft}) R_{f,t+1} \mathbb{1}\{w_{pft}^{Passive} \times w_{pft} \geq 0\} \tag{18}$$

$$R_{t+1}^{opposite} = \sum_{f \in \Omega_t} \sum_{p \in \Omega_{ft}} (w_{pft}^{Passive} - w_{pft}) R_{f,t+1} \mathbb{1}\{w_{pft}^{Passive} \times w_{pft} < 0\} \tag{19}$$

where $w_{pft}^{Passive}$ is the weight of fund $f$ in plan $p$ from the passive strategy, and $w_{pft}$ is the one from the observed flows.

Figure 3 shows the decomposition in performance between the observed flows and the passive strategy with 80% allocated in equity.[12] The entire performance difference can be explained by the first component, where the hypothetical flows and observed flows have opposite signs. Our results suggest that unsophisticated investors are better off avoiding active investing and simply passively investing in their 401(k) plans.

We also examine the long-term performance difference between the two strategies. In Panel (b) of Figure 2, we allow flows to accumulate in later years till the end of our sample period as follows:

$$Saving_{pft} = Flow_{pft} + Saving_{pf,t-1}(1 + R_{ft}), \text{ for } t \geq 2, \tag{20}$$

$$Saving_{pf1} = Flow_{pf1}, \tag{21}$$

where $Flow_{pft}$ is the flows to fund $f$ of plan $p$ in year $t$, and $Saving_{pft}$ is the total accu-

[12] The results are qualitatively similar when other cutoffs are used.

mulated savings for fund $f$ of plan $p$ in year $t$. Under our passive strategies, only the new flows ($Flow_{pft}$ for all $t$) will be reallocated according to the targeted parameter, while the accumulated savings ($Savings_{pft}$) from prior years will be reinvested without rebalancing again. The interpretation of this approach can be viewed as the cumulative performance of an employee who entered the workforce at the beginning of our sample period. Again, the passive strategy outperforms the observed flows regardless of the asset allocation role we use. In sum, the results suggest that participants can do well by passively investing in the plan rather than actively chasing funds with high past returns.

# 5 Conclusion

This paper examines how investors allocate wealth in 401(k) plans. Prior literature has debated whether investors either adjust returns for risk exposures (Barber et al. (2016), Berk and van Binsbergen (2016)) or rely on third-party signals (Ben-David et al. (2019), Evans and Sun (2018)). Instead of treating all market participants as a representative agent with unlimited attention who can browse through all funds and make investment decisions, we re-examine this question at the 401(k) plan level and focus on individual investors in the retirement market.

Using our hand-collected dataset on a large sample of defined contribution plans, we are the first to investigate the performance measures that individual investors pay attention to when they make investment decisions. We show that the majority of investors follow net-of-fee returns when choosing mutual funds. Moreover, we show there exist heterogeneous investment patterns across wealth levels within retirement market. Specifically, 17% of investors in our sample hold more than 61% of total assets, and their investment strategies are drastically different from the rest. These wealthy investors consider the systematic risk when selecting mutual funds. Lastly, we show that, by actively chasing funds with high past returns, these unwealthy investors leave substantial gains on the table. In particular, passively investing in all existing funds of the plans will outperform their return-chasing strategy by a large margin. The result is disconcerting, especially given that unwealthy investors typically do not invest outside their retirement account, do not have easy access to financial advisors, and rely heavily on this single account once they retire.

# References


Agnew, Julie, Pierluigi Balduzzi, and Annika Sunden, 2003, Portfolio choice and trading in a large 401 (k) plan, *American Economic Review* 93, 193–215.

Barber, Brad M., Xing Huang, and Terrance Odean, 2016, Which Factors Matter to Investors? Evidence from Mutual Fund Flows, *The Review of Financial Studies* 29, 2600–2642.

Barber, Brad M, and Terrance Odean, 2007, All that glitters: The effect of attention and news on the buying behavior of individual and institutional investors, *The Review of Financial Studies* 21, 785–818.

Barras, Laurent, Olivier Scaillet, and Russ Wermers, 2010, False discoveries in mutual fund performance: Measuring luck in estimated alphas, *The Journal of Finance* 65, 179–216.

Ben-David, Itzhak, Jiacui Li, Andrea Rossi, and Yang Song, 2019, What do mutual fund investors really care about?, *Fisher College of Business Working Paper* 005.

Benartzi, Shlomo, 2001, Excessive extrapolation and the allocation of 401 (k) accounts to company stock, *The Journal of Finance* 56, 1747–1764.

Benartzi, Shlomo, and Richard Thaler, 2007, Heuristics and biases in retirement savings behavior, *Journal of Economic Perspectives* 21, 81–104.

Benartzi, Shlomo, and Richard H Thaler, 2001, Naive diversification strategies in defined contribution saving plans, *American Economic Review* 91, 79–98.

Berk, Jonathan B, and Richard C Green, 2004, Mutual fund flows and performance in rational markets, *Journal of Political Economy* 112, 1269–1295.

Berk, Jonathan B, and Jules H van Binsbergen, 2016, Assessing asset pricing models using revealed preference, *Journal of Financial Economics* 119, 1–23.

Blume, Marshall E., 1998, An anatomy of morningstar ratings, *Financial Analysts Journal* 54, 19–27.

Bollerslev, Tim, 1986, Generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics* 31, 307–327.

Bosler, Canyon, Nicolas Petrosky-Nadeau, et al., 2016, Job-to-job transitions in an evolving labor market, *FRBSF Economic Letter* 34.

Calvet, Laurent E, John Y Campbell, and Paolo Sodini, 2007, Down or out: Assessing the welfare costs of household investment mistakes, *Journal of Political Economy* 115, 707–747.

Calvet, Laurent E, John Y Campbell, and Paolo Sodini, 2009, Measuring the financial sophistication of households, *American Economic Review* 99, 393–98.

Campanale, Claudio, 2007, Increasing returns to savings and wealth inequality, *Review of Economic Dynamics* 10, 646–675.

Carhart, Mark M, 1997, On persistence in mutual fund performance, *The Journal of Finance* 52, 57–82.

Cici, Gjergji, and Scott Gibson, 2012, The performance of corporate bond mutual funds: Evidence based on security-levelholdings, *The Journal of Financial and Quantitative Analysis* 47, 159–178.

Elton, Edwin J, Martin J Gruber, and Christopher R Blake, 1995, Fundamental economic variables, expected returns, and bond fund performance, *The Journal of Finance* 50, 1229–1256.

Evans, Richard B, and Yang Sun, 2018, Models or stars: The role of asset pricing models and heuristics in investor risk adjustment, *The Review of Financial Studies* .

Fama, Eugene F, and Kenneth R French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Lacetera, Nicola, Devin G Pope, and Justin R Sydnor, 2012, Heuristic thinking and limited attention in the car market, *American Economic Review* 102, 2206–36.

Lusardi, Annamaria, Pierre-Carl Michaud, and Olivia S Mitchell, 2017, Optimal financial knowledge and wealth inequality, *Journal of Political Economy* 125, 431–477.

Lusardi, Annamaria, and Olivia S Mitchell, 2014, The economic importance of financial literacy: Theory and evidence, *Journal of Economic Literature* 52, 5–44.

Ma, Linlin, Yuehua Tang, and Juan-Pedro Gomez, 2019, Portfolio manager compensation in the u.s. mutual fund industry, *The Journal of Finance* 74, 587–638.

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

Mitchell, Olivia S, and Stephen P Utkus, 2002, The role of company stock in defined contribution plans, Technical report, National Bureau of Economic Research.

Molloy, Raven, Riccardo Trezzi, Christopher L Smith, and Abigail Wozniak, 2016, Understanding declining fluidity in the us labor market, *Brookings Papers on Economic Activity* 2016, 183–259.

Pástor, L'uboš, Robert F Stambaugh, and Lucian A Taylor, 2017, Do funds make more when they trade more?, *The Journal of Finance* 72, 1483–1528.

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, 1779–1812.

Poterba, James M, 2003, Employer stock and 401 (k) plans, *American Economic Review* 93, 398–404.

Sialm, Clemens, Laura T Starks, and Hanjiang Zhang, 2015, Defined contribution pension plans: Sticky or discerning money?, *The Journal of Finance* 70, 805–838.

Sirri, Erik R, and Peter Tufano, 1998, Costly search and mutual fund flows, *The Journal of Finance* 53, 1589–1622.

Spataro, Luca, and Lorenzo Corsini, 2017, Endogenous financial literacy, saving and stock market participation, *FinanzArchiv/Public Finance Analysis* 73, 135–162.

Utkus, Stephen P, and Jean Young, 2011, Financial literacy and 401 (k) loans, *Financial literacy: Implications for retirement security and the financial marketplace* 59.

Figure 1
Distributions of 401(k) savings and the number of employees

The figure shows the distributions of 401(k) savings and the number of employees in our sample. In each year, we rank firms by the average savings per employee into percentiles, and calculate the cumulative fractions of the number of employees and total 401(k) savings, respectively. We then average the two cumulative fractions over the years for each percentile. The red vertical line indicates the cutoff of wealthy employees, who are in firms ranked in the top quartile (75th percentile).

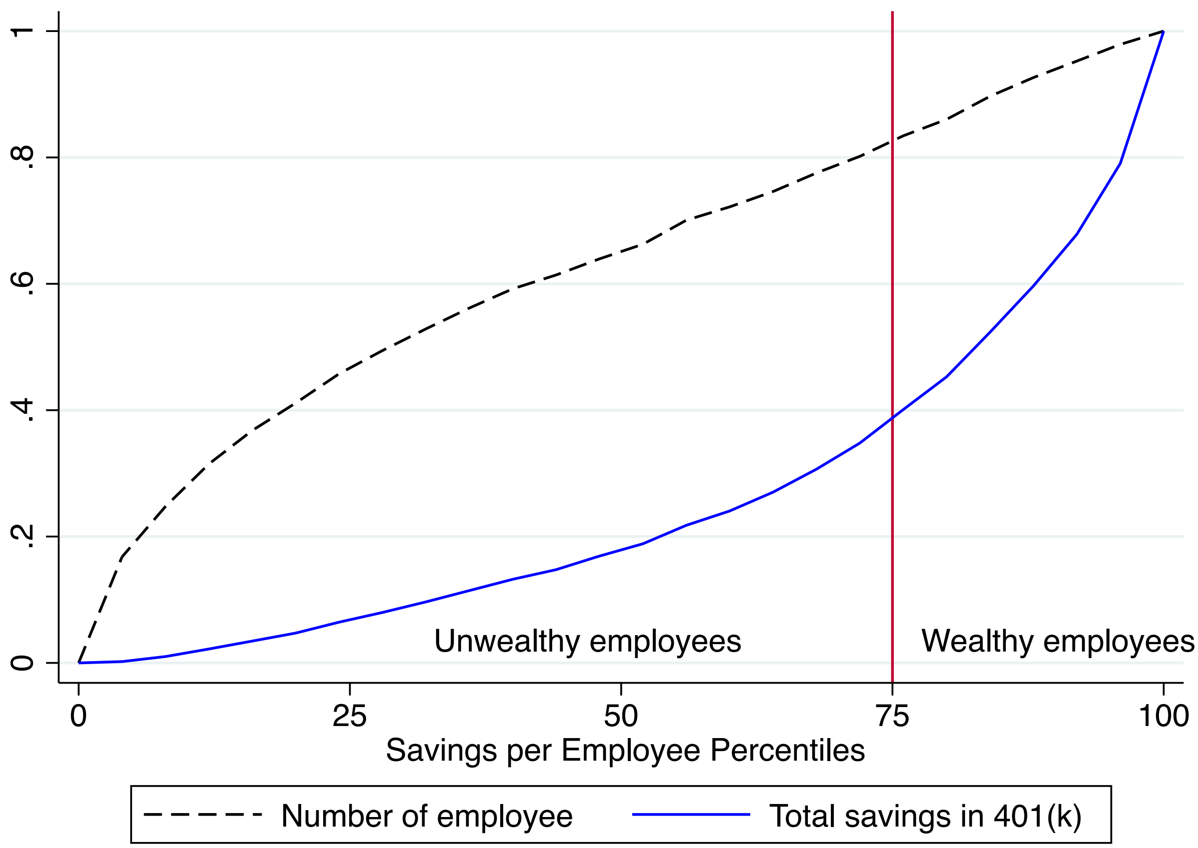


Figure 2

Performance of a passive strategy

The figure shows the performance of a passive investment strategy. Given an asset allocation rule that invests a certain fraction in equity, at the end of each year and for each plan, we redistribute the total dollar flows to the equity funds and bond funds in the plan. If the plan has more than one equity (bond) funds, flows are equally distributed among them. We consider three allocation cutoffs, 80%, 60%, and 40% in equity. Panel (a) studies the short-term performance of the observed and hypothetical flows over the next year, and we plot the cumulative performance during our sample period. Panel (b) studies the long-term performance by allowing flows to accumulate in later years till the end of our sample period.

(a) Short-term performance of flows without accumulation

(b) Long-term performance of flows with accumulation

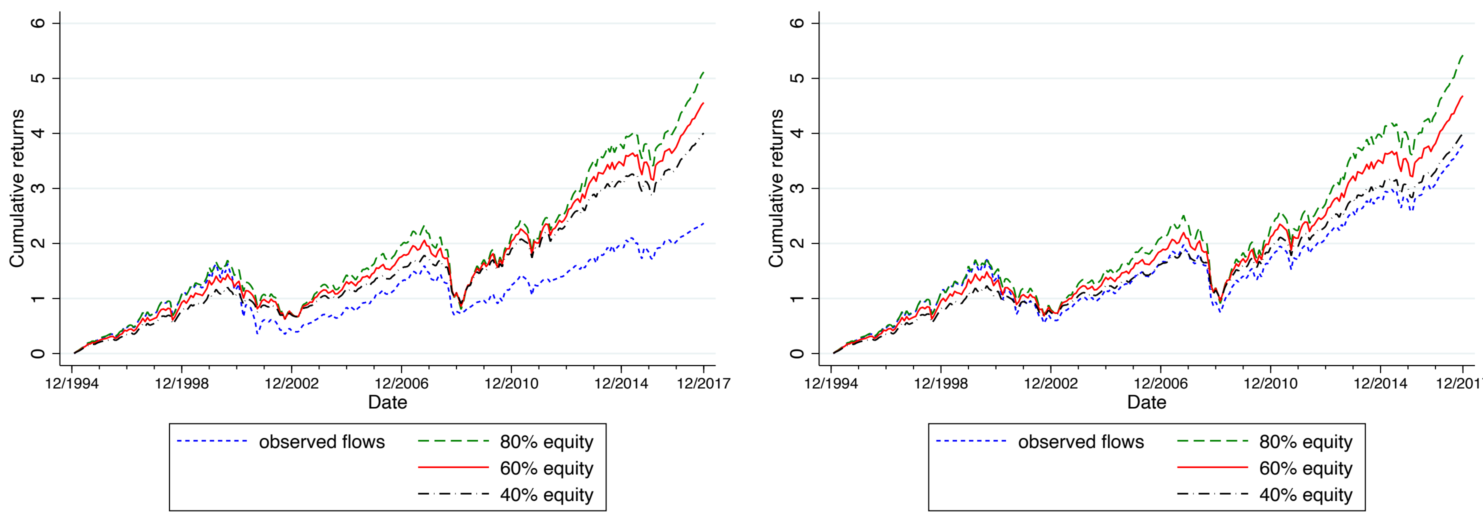


Figure 3
Decomposition of short-term performance difference

The figure shows the decomposition of the performance difference between the passive strategy with 80% allocated in equity and the observed flows into two components, depending on whether the observed flows and redistributed flows have the same signs. The dotted black line represents the cumulative performance difference between the two strategies. The red solid line represents the contribution of performance difference when observed flows and hypothetical flows have opposite signs. The blue dashed line represents the contribution of performance difference when observed flows and hypothetical flows have the same sign but potentially different magnitudes.

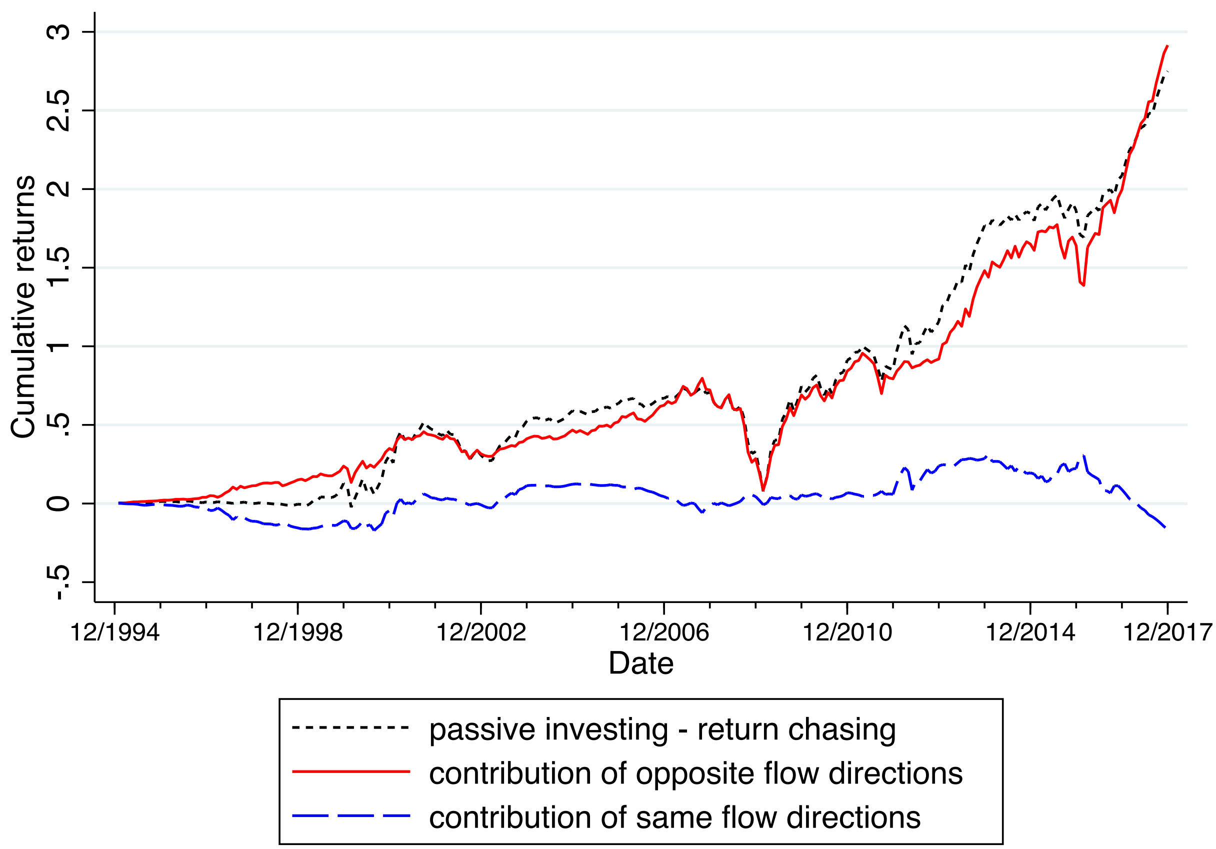


Table 1
Summary statistics

This table reports the descriptive statistics for our sample. Column 1 reports the number of firms with a 401(k) plan. Columns 2 and 3 show the average and median value of the plan size in millions. Column 4 reports the proportion of company stock held by the employees in their savings. Columns 5 and 6 show the total number of mutual funds invested in all plans and the average number of funds used in each plan. Columns 7 and 8 report the average fractions of capital investments in equity funds and bond funds to the total investment in mutual funds within a plan.

| Year | Number of firms | Plan size (millions) | | Company stock share (%) | Number of funds | Number of funds per plan | % investment in | |
|---|---|---|---|---|---|---|---|---|
| | | Mean | Median | | | | equity funds | bond funds |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| 1993 | 84 | 346 | 34 | 47.80 | 110 | 3 | 59.78 | 21.80 |
| 1994 | 154 | 224 | 34 | 46.02 | 167 | 3 | 59.25 | 25.20 |
| 1995 | 234 | 120 | 25 | 45.90 | 242 | 3 | 59.50 | 26.46 |
| 1996 | 309 | 163 | 25 | 46.23 | 325 | 4 | 61.87 | 25.26 |
| 1997 | 393 | 214 | 29 | 40.30 | 435 | 4 | 65.81 | 21.34 |
| 1998 | 466 | 214 | 28 | 35.13 | 550 | 5 | 66.85 | 20.15 |
| 1999 | 521 | 231 | 42 | 30.98 | 722 | 6 | 71.62 | 17.24 |
| 2000 | 598 | 320 | 37 | 31.19 | 858 | 7 | 72.90 | 16.89 |
| 2001 | 680 | 251 | 38 | 32.62 | 1,022 | 8 | 67.88 | 20.20 |
| 2002 | 738 | 196 | 35 | 32.40 | 1,142 | 9 | 62.83 | 24.37 |
| 2003 | 808 | 264 | 47 | 33.88 | 1,304 | 10 | 67.31 | 21.13 |
| 2004 | 828 | 268 | 59 | 32.62 | 1,390 | 11 | 70.26 | 18.02 |
| 2005 | 806 | 311 | 68 | 30.99 | 1,438 | 11 | 71.98 | 17.10 |
| 2006 | 774 | 388 | 80 | 30.53 | 1,427 | 11 | 72.69 | 16.36 |
| 2007 | 764 | 454 | 84 | 26.90 | 1,447 | 12 | 71.56 | 16.24 |
| 2008 | 737 | 267 | 59 | 25.35 | 1,450 | 13 | 62.39 | 23.55 |
| 2009 | 723 | 351 | 74 | 25.29 | 1,424 | 13 | 63.87 | 21.75 |
| 2010 | 678 | 434 | 93 | 24.71 | 1,425 | 14 | 64.99 | 20.20 |
| 2011 | 665 | 424 | 98 | 23.74 | 1,437 | 15 | 62.72 | 20.71 |
| 2012 | 637 | 484 | 125 | 23.36 | 1,426 | 15 | 62.12 | 20.57 |
| 2013 | 607 | 597 | 149 | 24.30 | 1,612 | 16 | 67.25 | 15.91 |
| 2014 | 582 | 615 | 159 | 23.68 | 1,671 | 17 | 67.90 | 14.78 |
| 2015 | 541 | 565 | 156 | 21.37 | 1,621 | 17 | 68.18 | 14.35 |
| 2016 | 473 | 596 | 164 | 24.16 | 1,163 | 14 | 68.31 | 14.15 |

Table 2
Flow-performance relation at the plan level

This table reports results from regressions of flows to fund $f$ of plan $p$ in year $t+1$ on fund performance, and fund characteristics in year $t$. The unit of observation is plan-fund-year. There are four different measures of fund performance: [1] $Net\ return_t$ is fund net-of-fee return; [2] $\alpha_t^{CAPM}$ is fund abnormal return which is adjusted for the CRSP value-weighted stock index (market) factor for equity and balanced funds and adjusted for the U.S. aggregate bond index for bond funds; [3] $\alpha_t^{4Factor}$ is fund abnormal return which uses the Carhart (1997) four-factor model for equity and balanced funds, and four-bond-factor for bond funds according to Ma et al. (2019), Cici and Gibson (2012), and Elton et al. (1995); [4] $MStar\ return_t$ is standardized Morningstar return within each investment category. Fund characteristics include expense ratio, turnover ratio, the logarithm of total fund net assets, fund return volatility, and investment style fixed effect. In all specifications, we have firm×year fixed effect. Standard errors are two-way clustered at the firm and year levels. The $t$-statistics are in brackets. *, **, and *** indicate significance at 10%, 5%, and 1% levels, respectively.

| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| | $Flow_{t+1}$ | $Flow_{t+1}$ | $Flow_{t+1}$ | $Flow_{t+1}$ | $Flow_{t+1}$ | $Flow_{t+1}$ | $Flow_{t+1}$ | $Flow_{t+1}$ |
| $Net\ return_t$ | 0.020*** | | | | 0.018** | 0.022** | 0.020*** | 0.019** |
| | [2.99] | | | | [2.48] | [2.65] | [2.82] | [2.59] |
| $\alpha_t^{CAPM}$ | | 0.021** | | | 0.005 | | | 0.013 |
| | | [2.63] | | | [0.70] | | | [1.50] |
| $\alpha_t^{4Factor}$ | | | 0.014** | | | -0.005 | | -0.014 |
| | | | [2.31] | | | [-0.72] | | [-1.66] |
| $MStar\ return_t$ | | | | 0.001** | | | 0.001 | 0.001 |
| | | | | [2.49] | | | [0.78] | [0.59] |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm×Year FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 118,168 | 118,168 | 118,168 | 118,168 | 118,168 | 118,168 | 118,168 | 118,168 |
| Adjusted $R^2$ | 0.42 | 0.42 | 0.42 | 0.42 | 0.42 | 0.42 | 0.42 | 0.42 |

Table 3
Pairwise horse-race

This table reports the pairwise horse-race between the net return model and the other competing models using the following regression:

$$Flow_{pf,t+1} = \sum_i \sum_j b_{ij} D_{ijpft} + \mathbf{X'}_{ft}\mathbf{c} + \mu_{pt} + \epsilon_{pf,t+1},$$

where $D_{ijpft}$ is a dummy variable that equals one if fund $f$ of firm $p$ in year $t$ is in tercile $i$ based on the fund net return and tercile $j$ based on other performance measures. To estimate the model, we exclude the dummy variable for $i = 2$ and $j = 2$. The matrix $\mathbf{X}_{ft}$ represents fund control variables, which are the same set of controls in Table 2. $\mu_{pt}$ is firm×year fixed effect. This table reports the sum of the differences between coefficients $b_{ij}$ and $b_{ji}$ for all $i$ and $j$ such that $i > j$:

$$\text{Sum} = \sum_{i>j}^{3} \sum_{j=1}^{2} (b_{i,j} - b_{j,i}).$$

Standard errors are two-way clustered at the firm and year levels. The $t$-statistics are in brackets. $^*$ and $^{**}$ indicate significance at the 10% and 5% levels, respectively.

| Winning Model | Net return | Net return | Net return |
|---|---|---|---|
| Losing Model | CAPM | 4Factor | Morningstar |
| Sum of coefficient differences | 0.006$^*$ | 0.009$^{**}$ | 0.004$^*$ |
| | [1.74] | [2.46] | [1.96] |

Table 4
Flows to top-ranked versus bottom-ranked funds

This table reports the average fund flows to the best and worst-performing funds. Specifically, for each firm in each year, we rank funds within the plan by various performance measures into terciles. The top tercile and bottom tercile contain the best and worst-performing funds. The panel *Positive Flows* shows the fraction of funds with positive flows. *Flows* and *Dollar Flows* are flows as a fraction of plan size and dollar flows, respectively. This table also reports the paired tests of the *Diff* (= Top - Bottom) between "Net return" terciles and other terciles. The $t$-statistics are in brackets. *** indicates significance at 1% levels.

| | Positive Flows (%) | | | Flows (%) | | | Dollar Flows (thousands) | | |
|---|---|---|---|---|---|---|---|---|---|
| | Top | Bottom | *Diff* | Top | Bottom | *Diff* | Top | Bottom | *Diff* |
| Net return | 67.00 | 61.54 | 5.47 | 1.74 | 1.02 | 0.71 | 880.33 | 315.31 | 565.02 |
| CAPM | 66.09 | 62.15 | 3.94 | 1.58 | 1.08 | 0.50 | 736.79 | 409.11 | 327.69 |
| 4Factor | 65.49 | 63.70 | 1.79 | 1.48 | 1.18 | 0.30 | 673.81 | 487.85 | 185.97 |
| Morningstar | 66.41 | 62.80 | 3.61 | 1.59 | 1.12 | 0.48 | 694.71 | 499.85 | 194.86 |
| The difference in *Diff* between Net return model and other models: | | | | | | | | | |
| - CAPM | | | 1.52*** | | | 0.21*** | | | 237.34*** |
| | | | [3.76] | | | [4.06] | | | [4.63] |
| - 4Factor | | | 3.68*** | | | 0.41*** | | | 379.06*** |
| | | | [8.14] | | | [7.30] | | | [6.57] |
| - Morningstar | | | 1.86*** | | | 0.23*** | | | 370.16*** |
| | | | [3.94] | | | [4.06] | | | [6.15] |

Table 5
Flow-performance relation at the aggregate level

This table studies the effect of menu change on flow-performance sensitivity in mutual funds. Columns (1) to (3) report results from fund-level regressions of aggregated flows to fund $f$ in year $t+1$ on fund performance and fund characteristics in year $t$. The dependent variables are total fund flows, sponsor flows, and participant flows, respectively. Columns (4) to (6) report results from plan-fund-level regressions. Column (4) reports plan-level sponsor flow regression. Columns (5) and (6) report plan-level regressions on two subsamples, one with menu change and one without. The control variables are the same as in Table 2. Fund performance and fund characteristics are described in Table 2. Standard errors are two-way clustered at the fund and year levels in columns (1) to (3) and at the firm and year levels in columns (4) to (6). The $t$-statistics are in brackets. *, **, and *** indicate significance at 10%, 5%, and 1% levels, respectively.

| | Aggregate level ($Flow_{f,t+1}$) | | | Micro level ($Flow_{pf,t+1}$) | | |
|---|---|---|---|---|---|---|
| | Total flow | Sponsor flow | Participant flow | Sponsor flow | Participant flow | |
| | | | | | Unchanged menu | Changed menu |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| *Net return$_t$* | 0.086 | 0.005 | 0.049 | 0.039 | 0.025** | 0.016** |
| | [0.81] | [0.16] | [0.74] | [1.57] | [2.52] | [2.50] |
| $\alpha_t^{CAPM}$ | 0.434** | 0.136*** | 0.231** | 0.032 | 0.019 | 0.009 |
| | [2.16] | [2.91] | [2.53] | [0.98] | [1.51] | [1.28] |
| $\alpha_t^{4Factor}$ | 0.168 | 0.100 | -0.110 | 0.007 | -0.016 | -0.012 |
| | [0.90] | [1.59] | [-1.65] | [0.27] | [-1.54] | [-1.69] |
| *MStar return$_t$* | 0.068*** | 0.041*** | 0.006 | 0.017*** | 0.001 | 0.000 |
| | [6.03] | [8.42] | [1.29] | [5.17] | [1.01] | [0.34] |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Fixed Effect | Year | Year | Year | Firm×Year | Firm×Year | Firm×Year |
| Observations | 23,333 | 23,333 | 23,333 | 18,380 | 47,842 | 70,326 |
| Adjusted $R^2$ | 0.04 | 0.06 | 0.03 | 0.26 | 0.49 | 0.38 |

Table 6
Heterogeneous flow-performance relation across employee savings

This table reports results from the plan-level regressions of flows to fund $f$ of plan $p$ in year $t+1$ on fund performance, plan and fund characteristics in year $t$, conditional on the rankings of per-employee savings. Fund performance, plan and fund characteristics are described in Table 2. In each year, we rank firms by the average 401(k) savings per employee into quartiles. In all specifications, we have firm×year fixed effect. Standard errors are two-way clustered at the firm and year levels. The $t$-statistics are shown in brackets. *, **, and *** indicate significance at 10%, 5%, and 1% levels, respectively.

| 401(k) savings/employee | Low | 2 | 3 | High |
|---|---|---|---|---|
| | $Flow_{t+1}$ | $Flow_{t+1}$ | $Flow_{t+1}$ | $Flow_{t+1}$ |
| $Net\ return_t$ | 0.022** | 0.026*** | 0.021** | 0.011 |
| | [2.38] | [2.97] | [2.77] | [1.54] |
| $\alpha_t^{CAPM}$ | 0.006 | 0.006 | 0.007 | 0.029*** |
| | [0.58] | [0.76] | [0.70] | [3.16] |
| $\alpha_t^{4Factor}$ | -0.013 | -0.013 | -0.010 | -0.021* |
| | [-1.25] | [-1.33] | [-1.24] | [-1.84] |
| $MStar\ return_t$ | -0.000 | 0.001 | 0.001 | 0.001 |
| | [-0.40] | [1.24] | [0.65] | [0.94] |
| Controls | Yes | Yes | Yes | Yes |
| Firm×Year FE | Yes | Yes | Yes | Yes |
| Observations | 26,210 | 27,115 | 30,974 | 33,112 |
| Adjusted $R^2$ | 0.49 | 0.39 | 0.36 | 0.33 |

Table 7
Plan and employee characteristics

This table reports the average values of various characteristics of 401(k) plans in each subsample. In each year, firms are sorted into quartiles by the firm's average 401(k) savings per employee. *Contribution cap* is the maximum amount that a firm is willing to contribute as a percentage of an employee's annual compensation. *Matching ratio* is the ratio of the employer's contribution to the employee's contribution. *401(k) contribution per employee* is the annual contribution to 401(k) plan per employee. The table also reports the difference in mean tests for all variables between the wealthy group (High) and the unwealthy group (bottom three quartiles). The standard errors are clustered at the year level, and *** indicates significance at 1% level.

| 401(k) savings/employee | Low | 2 | 3 | Wealthy (High) | Unwealthy |
|---|---|---|---|---|---|
| Savings per employee | 4,483 | 14,529 | 27,791 | 121,616 | 15,592*** |
| Plan size (in millions) | 128 | 170 | 304 | 840 | 201*** |
| Number of employees | 35,753 | 11,963 | 10,878 | 10,597 | 19,545*** |
| Firm age | 24.3 | 22.8 | 25.0 | 33.4 | 24.0*** |
| Firm matching contribution policy: | | | | | |
| - contribution cap (in %) | 5.37 | 5.45 | 5.46 | 5.54 | 5.43*** |
| - matching ratio | 0.46 | 0.52 | 0.54 | 0.63 | 0.51*** |
| 401(k) contribution per employee | 1,488 | 2,592 | 3,390 | 5,299 | 2,513*** |

Table 8
Flow gap between company stock and mutual funds

This table studies the sensitivity of the flow gap to the difference in performance between company stock and mutual funds in the plan. The dependent variable is the *Flow gap*, which is the difference in dollar flows between the company stock and all mutual funds in the plan, normalized by the plan size. The independent variable $\mathbb{1}\{\Delta PERF_{pt} > 0\}$, where $PERF = \{Net\ return, \alpha^{CAPM}, \alpha^{4Factor}\}$, is a dummy variable which is equal to one if the difference in performance between the firm's stock and aggregated funds in the firm's 401(k) plan is positive. In addition, for each year, we sort firms into quartiles by the 401(k) investment per employee. Column (2) shows the result of the flow gap sensitivity estimation for the *Wealthy* employees, who are in the highest average investment quartile. Column (1) presents the result for the *Unwealthy* ones, who are in the remaining groups. Firm characteristics include the logarithm of firm size, book-to-market ratio, the logarithm of the number of funds available in the plan, the logarithm of company stock and fund size in 401(k) plan. Fund characteristics include aggregated expense ratio, aggregated turnover ratio, the logarithm of total fund net assets. Standard errors are two-way clustered at the firm and year levels. The $t$-statistics are shown in brackets. *, **, and *** indicate significance at 10%, 5%, and 1% levels, respectively.

| | (1) | (2) |
|---|---|---|
| | $Flow\ gap_{t+1}$ | |
| | Unwealthy | Wealthy |
| $\mathbb{1}\{\Delta Net\ return_t > 0\}$ | 0.016* | 0.002 |
| | [1.75] | [0.19] |
| $\mathbb{1}\{\Delta\alpha_t^{CAPM} > 0\}$ | -0.006 | 0.027** |
| | [-0.51] | [2.22] |
| $\mathbb{1}\{\Delta\alpha_t^{4Factor} > 0\}$ | 0.000 | -0.014 |
| | [0.02] | [-1.35] |
| Controls | Yes | Yes |
| Firm FE | Yes | Yes |
| Year FE | Yes | Yes |
| Observations | 6,357 | 1,739 |
| Adjusted $R^2$ | 0.13 | 0.15 |

Table 9
Determinants of wealthy group

This table reports results from regressions of the dummy variable *Wealthy*, which is equal to one if employees are in the highest average savings quartile, on proxies of financial literacy and a set of plan, employee, and firm characteristics. There are three proxies for *Financial literacy*: in columns (1) and (4) *Loan* is the loans to participants scaled by 401(k) plan size; (2) and (5) *Loan Interest* is the loan interest scaled by the amount of loans; and (3) and (6) *Stock flow* is company stock flow to 401(k) plan. The higher the three measures, the lower the financial literacy. *Firm matching ratio* and *Firm contribution cap* are described in Table 7. *Firm age* is the logarithm of firm age. *Plan cumulative return* is value-weighted cumulative returns of all securities in 401(k) plan. *Employee contribution* is the logarithm of 401(k) contribution per employee. Firm characteristics include the logarithm of market capitalization, book-to-market ratio, debt-to-asset ratio, gross profitability ratio, and a dummy variable for whether the firm has defined benefit plan. Standard errors are two-way clustered at the firm and year levels. The $t$-statistics are shown in brackets. *, **, and *** indicate significance at 10%, 5%, and 1% levels, respectively.

| | (1) *Wealthy* | (2) *Wealthy* | (3) *Wealthy* | (4) *Wealthy* | (5) *Wealthy* | (6) *Wealthy* |
|---|---|---|---|---|---|---|
| *Financial literacy proxy* | Loan | Loan Interest | Stock flow | Loan | Loan Interest | Stock flow |
| | −3.95*** | −1.10*** | −0.28*** | −4.46*** | −1.18*** | −0.29*** |
| | [−7.39] | [−3.30] | [−7.36] | [−8.34] | [−3.70] | [−8.39] |
| *Firm matching ratio* | 0.14*** | 0.16*** | 0.16*** | 0.14*** | 0.15*** | 0.15*** |
| | [4.87] | [5.46] | [5.60] | [4.87] | [5.47] | [5.38] |
| *Firm contribution cap* | 0.67 | 0.40 | 0.37 | −0.39 | −0.61 | −0.41 |
| | [0.80] | [0.47] | [0.45] | [−0.52] | [−0.80] | [−0.53] |
| *Firm age* | 0.10*** | 0.10*** | 0.10*** | 0.07*** | 0.07*** | 0.07*** |
| | [6.11] | [6.18] | [5.65] | [4.53] | [4.74] | [4.10] |
| *Plan cumulative return* | 0.01 | 0.01 | 0.01 | 0.01 | 0.01 | 0.00 |
| | [0.82] | [0.74] | [0.50] | [0.81] | [0.68] | [0.38] |
| *Employee contribution* | 0.16*** | 0.18*** | 0.16*** | 0.14*** | 0.16*** | 0.14*** |
| | [11.47] | [12.18] | [10.77] | [9.53] | [9.93] | [9.06] |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Fixed Effect | Year | Year | Year | Industry×Year | Industry×Year | Industry×Year |
| Observations | 7,051 | 7,051 | 7,153 | 7,051 | 7,051 | 7,153 |
| Adjusted $R^2$ | 0.26 | 0.24 | 0.24 | 0.30 | 0.29 | 0.28 |

Table 10
Financial literacy across Fama-French 12 industries

The table shows the distributions of Fama-French 12 industries in unwealthy and wealthy groups, respectively, and the difference in financial literacy between the two groups across industries. First, for firms classified in the unwealthy (wealthy) group, we tabulate the relative frequency of each industry. Second, for each industry, we test and report the difference in financial literacy between wealthy and unwealthy groups. We use three measures to proxy for financial literacy: (1) *Loan* is the loans to participants scaled by 401(k) plan size; (2) *Loan Interest* is the loan interest scaled by the amount of loans; and (3) *Stock flow* is company stock flow to 401(k) plan. The higher the three measures, the lower the financial literacy. *, **, and *** indicate significance at 10%, 5%, and 1% levels, respectively.

| | Distribution | | Financial literacy proxy | | |
|---|---|---|---|---|---|
| Industry | Unwealthy | Wealthy | Loan | Loan Interest | Stock flow |
| BusEq | 9.4 | 9.0 | −0.52*** | −0.67** | −3.11*** |
| Chems | 2.8 | 4.5 | −0.44*** | 0.15 | −4.85** |
| Durbl | 3.0 | 1.7 | −0.68* | 1.03 | −4.08* |
| Enrgy | 3.2 | 10.7 | −1.40*** | −0.89*** | −5.08*** |
| Hlth | 5.1 | 4.0 | −1.19*** | 0.61* | −4.98** |
| Manuf | 13.4 | 8.7 | −1.04*** | −0.15 | −1.46 |
| Money | 25.7 | 26.1 | −0.56*** | −0.86*** | −4.33*** |
| NoDur | 4.8 | 2.8 | −1.09*** | 0.67 | −2.8 |
| Shops | 13.4 | 0.8 | −1.85*** | −0.57 | −5.02 |
| Telcm | 3.8 | 1.7 | −0.57* | −0.12 | −8.59*** |
| Utils | 2.3 | 23.3 | −0.41*** | −0.94*** | −9.67*** |
| Other | 13.2 | 6.6 | −1.00*** | 0.17 | −3.09*** |

Table 11
Flow-performance sensitivity of index and active funds

The table reports the flow-performance sensitivities of index funds and active funds, separately for unwealthy and wealthy employees. We regress flows to fund $f$ of plan $p$ in year $t+1$ on four performance measures, an index fund dummy, and their interaction terms, controlling for firm×year fixed effects and a set of fund characteristics. The set of fund characteristics is the same as the one in Table 2. Columns (1) and (3) report the coefficient estimates of the four performance measures that capture the flow-performance sensitivity among active funds, while columns (2) and (4) report the sum of the coefficient estimate of the performance measure and its interaction term with the index fund dummy to capture the sensitivity among index funds. Standard errors are two-way clustered at the firm and year levels. The significant levels of index funds are calculated using the F-test. The p-values are in parentheses. *, **, and *** indicate significance at 10%, 5%, and 1% levels, respectively.

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| | $Flow_{t+1}$ | | | |
| | Unwealthy | | Wealthy | |
| | Active | Index | Active | Index |
| $Net\ return_t$ | 0.021** | 0.023*** | 0.010 | 0.009 |
| | (0.012) | (0.009) | (0.162) | (0.145) |
| $\alpha_t^{CAPM}$ | 0.009 | 0.010 | 0.032*** | 0.020 |
| | (0.354) | (0.484) | (0.003) | (0.268) |
| $\alpha_t^{4Factor}$ | −0.009 | −0.038** | −0.020 | −0.027 |
| | (0.284) | (0.033) | (0.112) | (0.145) |
| $MStar\ return_t$ | 0.000 | 0.001 | 0.000 | 0.001 |
| | (0.585) | (0.401) | (0.474) | (0.166) |
| Controls | Yes | | Yes | |
| Firm×Year FE | Yes | | Yes | |
| Observations | 84,299 | | 33,112 | |
| Adjusted $R^2$ | 0.43 | | 0.34 | |

# Appendix

## A Defined benefit plans and labor market mobility

There are two potential concerns about whether our average savings capture the wealth level. The first concern is that some firms may also offer defined benefit plans. If our classified unwealthy group is more likely to have defined benefit plans than the wealthy group, we may underestimate the wealth level for the unwealthy group and overestimate the wealth level for the wealthy group. The second concern is that some employees may keep 401(k) plans with the previous employer, and our test implicitly assumes employees always transfer their savings into the new plan when they move. In this section, we discuss these two concerns in detail.

### A.1 Defined benefit plans

To rule out the concern that firms in our unwealthy group may be more likely to offer defined benefit plans, we hand-collect data on firms' defined benefit plans and 401(k) plans from Form 5500. For each firm and year, we examine whether the firm has offered defined benefit plans to its employees, the fraction of assets in 401(k) plans, and the fraction of annual contribution to 401(k) plans. Table A6 presents these key statistics for wealthy and unwealthy groups.

At the extensive margin, the unwealthy group is less likely to have Defined Benefit Plans than the wealthy group. Specifically, 31.5% (44.1% ) of the unwealthy (wealthy) group offer Defined Benefit Plans to employees. At the intensive margin, assets and annual contribution to 401(k) plans represent the vast majority of employees' savings with the firm. Unconditionally, over 82% of the assets are in 401(k) plans, and over 90% of the annual contribution is invested in 401(k) plans. Moreover, the unwealthy group has an even larger proportion of assets and contribution in 401(k) plans. Therefore, our wealthy classification does not underestimate the savings of the unwealthy group relative to the wealthy one.

## A.2 Labor market mobility

One could be concerned about whether the average savings in a plan capture the wealth level due to labor market mobility. In the US, when an employee switches jobs from one employer to another, she can either keep her 401(k) with the previous employer or transfer it to the new plan with the current employer. Our test above implicitly assumes employees always transfer their savings into the new plan when they move. As documented in Molloy, Trezzi, Smith, and Wozniak (2016), the US job mobility rate has declined over the past several decades to around 10% in the 2010s, and interstate migration only accounts for 2%. Given the stylized fact, we conduct the following robustness check to argue job mobility channel cannot explain the heterogeneous return-chasing pattern across wealth levels. To account for the fact that some employees may not transfer their savings to the new employer, we adjust our average savings per employee measure as follows:

$$\textit{Savings per employee}_{pt} = \frac{P_{pt} + \phi N_{pt} \textit{New Employee Saving}_{pt}}{N_{pt}}, \tag{A.1}$$

$$\textit{New Employee Saving}_{pt} = \frac{\sum_{c \in \Psi^X_{pt}} P_{ct}}{\sum_{c \in \Psi^X_{pt}} N_{ct}}, \tag{A.2}$$

where $N_{pt}$ is the number of firm $p$'s employees at year $t$; $P_{pt}$ is the total 401(k) investment value; $\phi$ is the fraction of employees who move to firm $p$ and do not transfer their funds from their old plan directly into the new employer's 401(k) plan. According to the job mobility rate and to be conservative, we set $\phi$ to be 15%. $\Psi_{pt}^{X}$ is the set of all firms that are located within $X$ miles of firm $p$, and we assume that employee mobility is within $X = 200$ miles of the current firm. Our estimation is not sensitive to the choice of $\phi$ or $X$.[13] Note that the second term of *Savings per employee* captures the savings new employees would have brought to the firm. We then sort firms into quartiles by the new *Savings per employee* into quartiles, and the results do not change, as shown in Appendix Table A7.

# B What explains 401(k) savings?

We define the total savings in the 401(k) plan of each employee up to time $t$ as follows (see Appendix C for detailed derivation):

$$Saving_{1\to t} = Salary_t \frac{(1+g)^t - (1+r)^t}{(g-r)(1+r)^{t-1}} \left[s + min\{f_{cap}, s \times f_{rate}\}\right] \tag{B.1}$$

where $Salary_t$ is the employee's salary at time $t$, $r$ is the annual growth rate of the salary, $g$ is the annual growth rate of plan assets, and $s$ is the employee's deferral rate. The firm's matching contribution cap, $f_{cap}$, is the maximum amount that a firm is willing to contribute, and it is typically represented as a fraction of an employee's annual compensation. The firm matching ratio, $f_{rate}$, is the ratio of the firm's contribution to the employee's contribution. For example, a matching ratio of 0.5 means that a firm will contribute $0.50 for every dollar an employee contributes to the plan. The salary growth rate $r$ is set to the average inflation rate from 1993 to 2016.

From equation (B.1), there are five factors that could potentially contribute to the savings

[13] We consider a wide range of the two parameters: $X$ can range from 50 to 1000 miles, and $\phi$ can range from 5% to 50%.

gap between the wealthy and unwealthy groups: the firm's matching contribution policy ($f_{cap}$ and $f_{rate}$), the employee's tenure ($t$), the annual growth rate of the plan's assets ($g$), the employee's salary, and the employee's deferral rate ($s$). We then show that the first three factors only explain a small proportion of the savings gap, leaving a large part to be explained by either the wage or the deferral rate.

First, we show that the firm's matching policy only explains 1% of the savings gap. Table 7 shows that firms will match up to 5.54% of employee compensation for the wealthy group and 5.43% for the unwealthy group. The difference in contribution cap is statistically significant but economically small. In terms of contribution ratio, employees from the wealthy group receive from their employers an extra \$0.63 for every dollar they contribute to the plan compared to \$0.51 for employees from the unwealthy group. Even though the difference of \$0.12 is statistically significant, it does not explain much of the variation in the savings gap. For every dollar employees invest in the plan, those in the wealthy group will have 8% more capital compared to those in the unwealthy group. The 8% additional capital accounts for only 1% of the difference in savings of \$106,024 between the wealthy group and the unwealthy group.

Second, the massive savings gap cannot be entirely attributed to the long job tenure of the wealthy group. We use the difference in firm age as a proxy for the difference in job tenure. Firm age is defined as the time since we have firm data from either CRSP or Compustat datasets. The average difference in firm age between the two groups is about 10 years, and the difference in savings is \$106,024. If the difference in firm ages fully explains the savings gap, employees in both groups have to earn an annual wage of \$123,998.[14] However,

[14]Using equation (B.1) with the plan size growth rate is the average values of plan's historical investment returns and the employee deferral rate of 7%, from Vanguard defined contribution annual report, and accounts for non-participate rate of 33% from Department of Labor.

the income per capita in the U.S. is \$46,550 in 2016,[15] suggesting that the savings gap is unlikely to be due to the difference in job tenure. Furthermore, the 10-year gap in firm age is likely to overestimate the actual tenure gap, making the tenure gap even more unlikely to explain the savings gap.[16] In Appendix D, we also infer employees' age difference from the name of the target-date fund, which suggests the projected year of retirement. The tenure gap using the inferred age is only 1.6 years.

To quantify the contribution of the tenure gap to the savings gap, we use equation (B.1) and assume that the employee salary is equal to the income per capita in the U.S. (\$46,550). We use the group averages of the firm's matching contribution caps $f_{cap}$ and the firm's matching ratios $f_{rate}$ from Table 7. The growth rates of assets in the plans are set to the average performance of plans from 1993 to 2016, which are 6.4% and 5.2% for the wealthy and unwealthy groups, respectively. If we consider an extreme case where the difference in the tenure gap is the difference in the firm's age, the tenure gap contributes 20% of the difference in savings, and the plan's investment return difference contributes another 4%.

Therefore, the difference in savings is mainly driven by the difference in employees' deferral rate and their annual compensation. To determine the contributions of these two effects, we estimate the ratios of the deferral rate $s$ and the $Salary_t$ between the wealthy and unwealthy groups using equation (B.1) with the tenure gap of 9.4 years. Figure A2 illustrates the combination of these two ratios that help explain the remaining 75% of the savings

[15] Data is from The Census Bureau https://www.census.gov/data/tables/time-series/demo/income-poverty/cps-pinc/pinc-01.html.

[16] With a subsample where we have data on the firm's founding date, the difference in firm age is 4 years. In addition, according to the Department of Labor, the median tenure of workers with ages from 55 to 64 is 10.3, whereas that of younger workers with ages from 25 to 34 is 2.8 as of 2016. Data is from https://www.bls.gov/news.release/tenure.t01.htm. Consider an extreme case, in which the firm in the wealthy group is full of workers with ages between 55 and 64, and the firm in the lowest savings group is full of workers with ages between 25 and 34. Then the tenure gap between the two extreme firms is only 7.5, which is smaller than the firm age difference.

gap. For example, if employees from these two groups have the same salary and unwealthy employees save 3% per year,[17] wealthy employees would need to save around 12% of their income so that the difference in savings can be justified. On the other hand, if the deferral rate is the same, the median wealthy employees would earn annually around $139,650, assuming unwealthy employees earned an average income of $46,550 in the U.S. Therefore, either a high deferral rate or a high salary, or both, play a primary role in explaining the savings gap.

# C Determinants of savings

If an employee contributes a deferral rate of $s$ to the 401(k) plan each year, the firm will provide a matching contribution of $min\{f_{cap}, s \times f_{rate}\}$. The total contribution with respect to the employee's compensation at each year to the plan is $Saving\ rate = s + min\{f_{cap}, s \times f_{rate}\}$. Denote $Salary_1$ be a salary that the employee receives when he joins the firm, then the capital that employee has in the plan in the second year derives from two sources:

1. Plan growth:
$$PG_2 = Saving_{1\to 1} \times G = Salary_1 \times Saving\ rate \times G, \tag{C.1}$$

2. New contribution:
$$NC_2 = Salary_2 \times Saving\ rate = Salary_1 \times R \times Saving\ rate, \tag{C.2}$$

$\Rightarrow$ Total saving:
$$Saving_{1\to 2} = Salary_1 \times Saving\ rate \times \big(G + R\big) = Salary_1 \times Saving\ rate \times \frac{G^2 - R^2}{G - R}, \tag{C.3}$$

[17] Vanguard reports that the median employers have 401(k) plans that automatically enroll their employees at a 3% deferral rate.

where $G = 1 + g$ and $R = 1 + r$, and the total savings in year $t$ is:

$$Saving_{1 \to t} = Salary_1 \times Saving\ rate \times \frac{G^t - R^t}{G - R}. \tag{C.4}$$

# D Inferring age from target-date funds

We infer employees' ages from the name of the target-date Funds to estimate the employee's typical year of retirement. We use the difference in employees' ages as a proxy for the difference in job tenure. Our inferred age difference between the wealthy and unwealthy groups is only 1 years. Specifically, let a target-date fund $f$ have a targeted utilization year of $T_f$. We assume that if employees in year $t$ invest in this fund, they will retire around year $T_f$. Therefore, the average age of employees who invest in this target-date fund $f$ in year $t$ is:

$$Age^{TA}_{pft} = 65 - (T_f - t) \tag{D.1}$$

Panel 1 in Table A8 shows that the difference in employee age between the highest and lowest savings groups is 1 year if only target-date funds are used. Next, we predict the average age of employees (denote $Age^{NT}_{pft}$) of firm $p$ who invest in non-target-date fund $f$ in year $t$ using the model as follows:

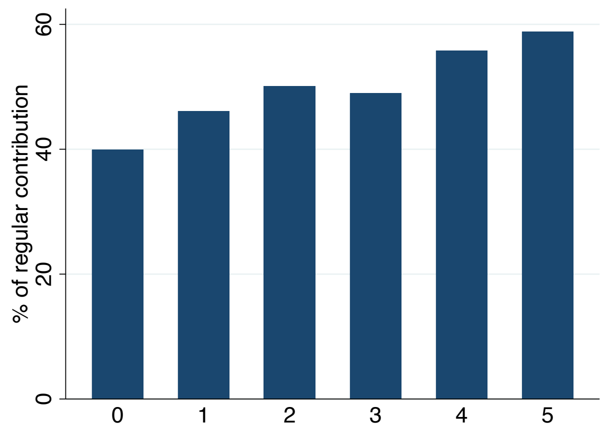
$$Age^{O}_{pft} = \Phi\Big(\mathbf{N'}_{pft}\boldsymbol{\beta} + \lambda_p + \kappa_t\Big) + \zeta_{pft}, \tag{D.2}$$

$$Age^{O}_{pft} = \frac{Age^{K}_{pft} - 22}{65 - 22} \in (0, 1), \tag{D.3}$$

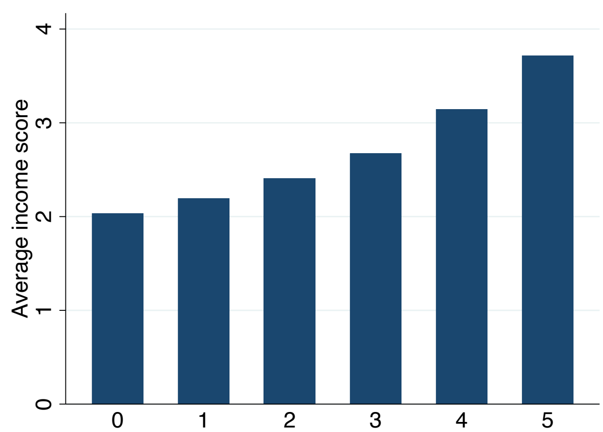
and $K = \{TA, NT\}$; $\Phi(.)$ is the cumulative distribution function of the standard normal distribution, and $\zeta_{pft}$ is normally distributed with a zero mean. $\mathbf{N}_{pft}$ represents fund control variables which are the fraction of investment in plan assets, the logarithm of total fund net assets, expense ratio, turnover ratio, the fund's return volatility, and the fund's time-varying loading on market, size, and value factors. The assumption is that employees with a similar age invest in funds with similar characteristics and risk exposure. Panel 2 in Table A8 shows that the model has high predictive power, and the pseudo $R^2$ is 94%. Accounting for predicted ages, the difference in employee age between the wealthy and unwealthy groups is 1.6 years, which is clearly not a primary contribution to the difference in savings.

Figure A1
Financial literacy, regular contribution, and compensation

The figure shows that financial literacy is correlated with employee compensation and regular 401(k) contribution. The data is from surveys collected by the Foundation's National Financial Capability Study. Panel (a) shows the relation between the finance test score and the regularity of the contribution to the retirement plans. Panel (b) exhibits the relation between the finance test score and the income score. The income score ranges from 1, for those who have income less than \$25,000, to 6, for those who earn at least \$150,000.




(a) Percentage of regular contribution to retirement plans by finance test score.




(b) Average income score by finance test score.

Figure A2
Contributions of salary and deferral rate to the savings gap

The figure shows how the combination of the deferral rate ratio and the salary ratio between the wealthy and unwealthy groups can explain their differences in savings, after accounting for the differences in the firm's matching contribution policy, the employee tenure, and the plan's investment returns.

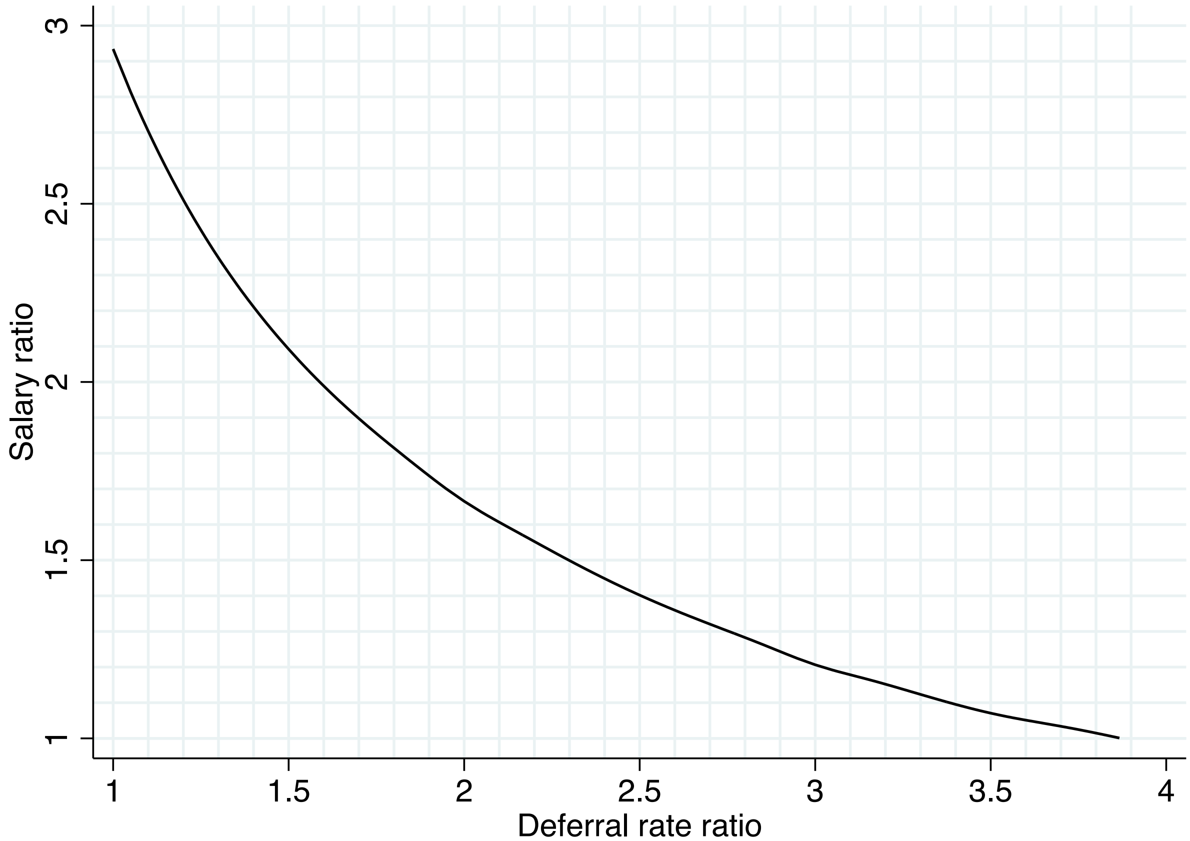


Table A1

Flow-performance relation at the plan level: at 3-year performance horizon

This table reports results from regressions of flows to fund $f$ of plan $p$ in year $t+1$ on fund performance, and fund characteristics in year $t$. The model is the same as that in Table 2. $Net\ return_{[t-2,t-1]}$, $\alpha^{CAPM}_{[t-2,t-1]}$, and $\alpha^{4Factor}_{[t-2,t-1]}$ are fund performances from year $t-2$ to $t-1$. In all specifications, we have firm×year fixed effect. Standard errors are two-way clustered at the firm and year levels. The $t$-statistics are in brackets. *, **, and *** indicate significance at 10%, 5%, and 1% levels, respectively.

| | (1) $Flow_{t+1}$ | (2) $Flow_{t+1}$ | (3) $Flow_{t+1}$ | (4) $Flow_{t+1}$ |
|---|---|---|---|---|
| $Net\ return_t$ | 0.021*** | | | 0.021** |
| | [2.99] | | | [2.56] |
| $Net\ return_{[t-2,t-1]}$ | 0.005 | | | 0.005 |
| | [1.18] | | | [0.59] |
| $\alpha^{CAPM}_t$ | | 0.021** | | 0.012 |
| | | [2.57] | | [1.39] |
| $\alpha^{CAPM}_{[t-2,t-1]}$ | | 0.000 | | -0.013 |
| | | [0.06] | | [-0.90] |
| $\alpha^{4Factor}_t$ | | | 0.014** | -0.014 |
| | | | [2.22] | [-1.68] |
| $\alpha^{4Factor}_{[t-2,t-1]}$ | | | 0.008 | 0.017 |
| | | | [1.18] | [1.56] |
| $MStar\ return_t$ | | | | -0.001 |
| | | | | [-0.09] |
| Controls | Yes | Yes | Yes | Yes |
| Firm×Year FE | Yes | Yes | Yes | Yes |
| Observations | 118,168 | 117,785 | 117,785 | 117,785 |
| Adjusted $R^2$ | 0.42 | 0.42 | 0.42 | 0.42 |

Table A2

Flow-performance relation of wealthy participants — Omitting industries one-by-one

The table shows the flow-performance relation of wealthy participants is not driven by a specific industry. The regression specification is identical to the one in column (4) of Table 6. We drop each industry one-by-one and report the regression result. Standard errors are two-way clustered at the firm and year levels. The $t$-statistics are shown in brackets. *, **, and *** indicate significance at 10%, 5%, and 1% levels, respectively.

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| | $Flow_{t+1}$ | $Flow_{t+1}$ | $Flow_{t+1}$ | $Flow_{t+1}$ | $Flow_{t+1}$ | $Flow_{t+1}$ |
| $Net\ return_t$ | 0.011 | 0.010 | 0.011 | 0.012 | 0.011 | 0.011 |
| | [1.58] | [1.42] | [1.54] | [1.56] | [1.56] | [1.60] |
| $\alpha_t^{CAPM}$ | 0.030*** | 0.029*** | 0.029*** | 0.029*** | 0.029*** | 0.028*** |
| | [3.13] | [3.09] | [3.19] | [3.17] | [3.26] | [3.12] |
| $\alpha_t^{4Factor}$ | -0.024* | -0.020 | -0.019* | -0.022* | -0.022* | -0.021* |
| | [-1.82] | [-1.63] | [-1.85] | [-1.85] | [-1.90] | [-1.79] |
| $MStar\ return_t$ | 0.000 | 0.001 | 0.000 | 0.000 | 0.001 | 0.001 |
| | [0.68] | [1.04] | [0.86] | [1.01] | [0.93] | [1.05] |
| Omitted Industry | BusEq | Chems | Durbl | Enrgy | Hlth | Manuf |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm×Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 29,446 | 31,792 | 32,481 | 29,168 | 31,916 | 30,699 |
| Adjusted $R^2$ | 0.34 | 0.34 | 0.34 | 0.33 | 0.32 | 0.33 |

| | (7) | (8) | (9) | (10) | (11) | (12) |
|---|---|---|---|---|---|---|
| | $Flow_{t+1}$ | $Flow_{t+1}$ | $Flow_{t+1}$ | $Flow_{t+1}$ | $Flow_{t+1}$ | $Flow_{t+1}$ |
| $Net\ return_t$ | 0.013 | 0.011 | 0.011 | 0.011 | 0.008 | 0.012 |
| | [1.65] | [1.59] | [1.53] | [1.52] | [1.24] | [1.62] |
| $\alpha_t^{CAPM}$ | 0.033*** | 0.029*** | 0.030*** | 0.030*** | 0.028*** | 0.031** |
| | [2.91] | [3.06] | [3.16] | [3.17] | [4.03] | [2.75] |
| $\alpha_t^{4Factor}$ | -0.027* | -0.021* | -0.022* | -0.022* | -0.018* | -0.023* |
| | [-1.94] | [-1.76] | [-1.85] | [-1.84] | [-1.97] | [-1.84] |
| $MStar\ return_t$ | 0.000 | 0.000 | 0.001 | 0.001 | 0.001 | 0.000 |
| | [0.58] | [0.90] | [0.95] | [0.98] | [1.43] | [0.80] |
| Omitted Industry | Money | NoDur | Shops | Telcm | Utils | Other |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Firm×Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 24,477 | 32,170 | 32,839 | 32,701 | 25,855 | 30,688 |
| Adjusted $R^2$ | 0.35 | 0.33 | 0.34 | 0.34 | 0.35 | 0.34 |

Table A3
Robustness of Table 9: subsample analysis on firms with and without DB plans

This table shows the robustness of results in Table 9 in two subsamples, firms with and without DB plans. Standard errors are two-way clustered at the firm and year levels. *, **, and *** indicate significance at 10%, 5%, and 1% levels, respectively.

| | Without DB | | | With DB | | |
|---|---|---|---|---|---|---|
| | (1) *Wealthy* | (2) *Wealthy* | (3) *Wealthy* | (4) *Wealthy* | (5) *Wealthy* | (6) *Wealthy* |
| *Financial literacy proxy* | Loan | Loan Interest | Stock flow | Loan | Loan Interest | Stock flow |
| | −4.23*** | −0.75* | −0.32*** | −5.20*** | −1.31** | −0.27*** |
| | [−6.38] | [−1.99] | [−6.34] | [−5.87] | [−2.66] | [−5.91] |
| *Firm matching ratio* | 0.14*** | 0.17*** | 0.18*** | 0.15*** | 0.15*** | 0.14*** |
| | [3.74] | [4.39] | [4.65] | [3.28] | [3.29] | [3.21] |
| *Firm contribution cap* | −0.69 | −0.98 | −0.93 | 0.33 | 0.04 | 0.74 |
| | [−0.85] | [−1.22] | [−1.13] | [0.20] | [0.02] | [0.49] |
| *Firm age* | 0.07*** | 0.07*** | 0.07*** | 0.07*** | 0.08*** | 0.06** |
| | [3.44] | [3.48] | [3.31] | [3.21] | [3.41] | [2.57] |
| *Plan cumulative return* | 0.03** | 0.03 | 0.02 | −0.00 | −0.00 | −0.00 |
| | [2.12] | [1.74] | [1.55] | [−0.19] | [−0.21] | [−0.40] |
| *Employee contribution* | 0.10*** | 0.12*** | 0.10*** | 0.18*** | 0.19*** | 0.19*** |
| | [8.05] | [8.46] | [7.79] | [6.88] | [7.28] | [6.48] |
| Controls | Yes | Yes | Yes | Yes | Yes | Yes |
| Industry×Year FE | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 3,305 | 3,305 | 3,415 | 3,735 | 3,735 | 3,726 |
| Adjusted $R^2$ | 0.27 | 0.24 | 0.25 | 0.31 | 0.30 | 0.29 |

Table A4
Robustness of Table 9: industry × state fixed effects

This table shows the robustness of results in Table 9 using industry × state fixed effects and year fixed effects. Standard errors are two-way clustered at the firm and year levels. The $t$-statistics are shown in brackets. *, **, and *** indicate significance at 10%, 5%, and 1% levels, respectively.

| | (1) *Wealthy* | (2) *Wealthy* | (3) *Wealthy* |
|---|---|---|---|
| *Financial literacy proxy* | Loan | Loan Interest | Stock flow |
| | $-4.81^{***}$ | $-1.31^{***}$ | $-0.25^{***}$ |
| | [−8.49] | [−4.83] | [−8.13] |
| *Firm matching ratio* | $0.13^{***}$ | $0.15^{***}$ | $0.15^{***}$ |
| | [4.52] | [5.00] | [4.86] |
| *Firm contribution cap* | 0.47 | 0.18 | 0.49 |
| | [0.51] | [0.20] | [0.56] |
| *Firm age* | $0.08^{***}$ | $0.07^{***}$ | $0.07^{***}$ |
| | [4.66] | [4.59] | [4.13] |
| *Plan cumulative return* | 0.01 | 0.00 | 0.00 |
| | [0.57] | [0.50] | [0.39] |
| *Employee contribution* | $0.13^{***}$ | $0.14^{***}$ | $0.13^{***}$ |
| | [7.93] | [8.10] | [8.19] |
| Controls | Yes | Yes | Yes |
| Year FE | Yes | Yes | Yes |
| Industry×State FE | Yes | Yes | Yes |
| Observations | 7,038 | 7,038 | 7,142 |
| Adjusted $R^2$ | 0.41 | 0.40 | 0.41 |

Table A5
Robustness of Table 9: using inferred employee age from target date funds in the plan

This table shows the robustness of results in Table 9 using the inferred employee age from target date funds in the plan, instead of firm age. Standard errors are two-way clustered at the firm and year levels. The *t*-statistics are shown in brackets. *, **, and *** indicate significance at 10%, 5%, and 1% levels, respectively.

| | (1) *Wealthy* | (2) *Wealthy* | (3) *Wealthy* |
|---|---|---|---|
| *Financial literacy proxy* | Loan | Loan Interest | Stock flow |
| | −4.61*** | −1.35*** | −0.34*** |
| | [−8.33] | [−3.95] | [−8.85] |
| *Firm matching ratio* | 0.13*** | 0.15*** | 0.15*** |
| | [4.37] | [4.92] | [4.72] |
| *Firm contribution cap* | −0.54 | −0.76 | −0.56 |
| | [−0.70] | [−0.98] | [−0.72] |
| *Employee age* | 0.17** | 0.18*** | 0.15** |
| | [2.92] | [3.15] | [2.66] |
| *Plan cumulative return* | 0.01 | 0.01 | 0.01 |
| | [0.96] | [0.83] | [0.47] |
| *Employee contribution* | 0.14*** | 0.16*** | 0.13*** |
| | [9.03] | [9.49] | [8.42] |
| Controls | Yes | Yes | Yes |
| Industry×Year FE | Yes | Yes | Yes |
| Observations | 6,867 | 6,867 | 6,947 |
| Adjusted $R^2$ | 0.30 | 0.28 | 0.28 |

Table A6
Defined benefit plan

The table reports: [1] percentage of firms have both defined benefit (DB) and 401(k) plans, [2] percentage of total contribution in a given year to 401(k) plan, and [3] percentage of total assets in 401(k) plan ([2] and [3] are both relative to the sum of DB and 401(k) plan). The table also reports the difference in mean tests for all variables between the wealthy and unwealthy groups. The standard errors are clustered at the firm level. The $t$-statistics are shown in brackets, and *** indicates significance at 1% level.

| | Unwealthy (1) | Wealthy (2) | Difference (2) - (1) |
|---|---|---|---|
| % of firms with DB plans | 31.5 | 44.1 | 12.6*** |
| | | | [4.67] |
| % of contribution to 401(k) | 92.1 | 87.3 | -4.8*** |
| | | | [-4.34] |
| % of assets in 401(k) | 85.9 | 80.0 | -5.9*** |
| | | | [-4.08] |

Table A7
Heterogeneous flow-performance relation across employee savings: 200-miles mobility

This table reports results from the plan-level regressions of flows to fund $f$ of plan $p$ in year $t+1$ on fund performance, plan and fund characteristics in year $t$, conditional on the rankings of per-employee savings. The estimation is similar to that from Table 6. However, the average 401(k) savings per employee at each firm is adjusted to account for employees who join the firm but left their 401(k) savings at the firm they ended their employment earlier. We assume that these employees move from firms that are located within 200 miles of the current firm. In all specifications, we have firm×year fixed effect. Standard errors are two-way clustered at the firm and year levels. The $t$-statistics are shown in brackets. *, **, and *** indicate significance at 10%, 5%, and 1% levels, respectively.

| 401(k) savings/employee | Low | 2 | 3 | High |
|---|---|---|---|---|
| | $Flow_{t+1}$ | $Flow_{t+1}$ | $Flow_{t+1}$ | $Flow_{t+1}$ |
| $Net\ return_t$ | 0.020** | 0.027*** | 0.021** | 0.011 |
| | [2.38] | [2.89] | [2.67] | [1.57] |
| $\alpha_t^{CAPM}$ | 0.006 | 0.005 | 0.009 | 0.029*** |
| | [0.61] | [0.60] | [0.92] | [3.13] |
| $\alpha_t^{4Factor}$ | -0.009 | -0.019* | -0.009 | -0.022* |
| | [-0.94] | [-1.72] | [-1.09] | [-1.88] |
| $MStar\ return_t$ | -0.000 | 0.001 | 0.000 | 0.000 |
| | [-0.48] | [1.66] | [0.12] | [0.93] |
| Controls | Yes | Yes | Yes | Yes |
| Firm×Year FE | Yes | Yes | Yes | Yes |
| Observations | 26,187 | 26,997 | 30,546 | 32,884 |
| Adjusted $R^2$ | 0.49 | 0.39 | 0.38 | 0.33 |

Table A8
Employee age

Panel 1 reports the average ages of employees who invested in target-date funds. The table also reports the difference in mean tests between the wealthy group (High) and unwealthy group (Exclude High). *Fund weight* is the fraction of investment in fund $f$ within 401(k) plan $p$ in year $t$. $\beta_{MKT}$, $\beta_{SMB}$, and $\beta_{HML}$ are the fund's time-varying loading on market, size, and value factors. Panel 2 reports the coefficients from the employee age prediction. Standard errors are clustered at the year level. *, **, and *** indicate significance at 10%, 5%, and 1% levels, respectively.

**Panel 1:** Average age.

| 401(k) savings/employee | Low | 2 | 3 | High | Exclude High |
|---|---|---|---|---|---|
| Age from target date funds | 48.3 | 48.9 | 49.3 | 49.9 | 48.9*** |
| Observations (firm-year) | 545 | 675 | 846 | 942 | 2,080 |
| | | | | | |
| Age from target-date funds and prediction model | 43.6 | 45.4 | 45.8 | 46.6 | 45.0*** |
| Observations (firm-year) | 1,334 | 1,385 | 1,557 | 1,628 | 4,290 |

**Panel 2:** Age prediction.

| | Coefficients | $t$-stat |
|---|---|---|
| *Fund Weight* | 1.10*** | [5.64] |
| *Fund Weight*$^2$ | −2.56*** | [−4.61] |
| *Fund Size* | −0.42*** | [−3.95] |
| *Fund Size*$^2$ | 0.04*** | [6.54] |
| *Expense Ratio* | 1.25*** | [4.03] |
| *Expense Ratio*$^2$ | −0.89*** | [−3.93] |
| *Turnover Ratio* | 0.10* | [1.95] |
| *Fund Return Volatility* | −0.55 | [−0.09] |
| $\beta_{MKT}$ | −5.06*** | [−14.01] |
| $\beta_{SMB}$ | −0.94** | [−2.15] |
| $\beta_{HML}$ | −0.04 | [−0.08] |
| Firm FE | Yes | |
| Year FE | Yes | |
| Observations | 16,179 | |
| Pseudo $R^2$ | 0.94 | |