Barking Up The Wrong Tree: Return-chasing in Mutual Funds

Anh Tran and Pingle Wang*

January 9, 2020

This paper examines how investors allocate their savings at the micro-level. Using a hand-collected dataset consisting of firm-level investment decisions of employees in 401(k) plans, we characterize the return-chasing behavior of the median investor and demonstrate a lack of financial literacy among investors in the retirement market. Specifically, we show that only 17% of the population with a high level of financial sophistication hold 61% of the wealth and invest based on the CAPM alpha, whereas 83% of the population chase unadjusted returns and leave substantial money on the table.

Keywords: Mutual funds, defined contribution, fund flows, wealth inequality, financial literacy

*Tran: Department of Finance, University of Connecticut, One University Place, Stamford, CT 06901, email: anh.2.tran@uconn.edu. Wang: Simon Business School, University of Rochester, 4-339 Carol Simon Hall Rochester, NY 14627, email: pingle.wang@simon.rochester.edu

1 Introduction

How investors allocate their savings to mutual funds is one of the most important questions in the asset management literature. Recent studies suggest that investors value funds that deliver superior returns in excess of the market exposure (Barber et al. (2016), Berk and van Binsbergen (2016)). Since these studies only observed money flows at the fund level, their results characterize the investment decision of the representative agent in the economy. In this paper, we study the asset selection problems in employer-sponsored defined contribution (DC) plans, through which we observe the investment decisions of employees in each plan. In contrast to the flow-performance relation of the representative agent, our paper is the first to test the flow-performance relation at the micro-level directly. Our results suggest that, at the micro-level, investors chase funds with high unadjusted returns and pay little attention to systematic risks or third-party ratings (Ben-David et al. (2019)). However, flows respond to the CAPM alpha at the aggregated fund level, consistent with the findings in the existing literature. The seemingly inconsistent results between the micro-level and aggregated level can be explained by a wealth inequality channel. In particular, over 61% of the aggregated wealth is held by only 17% of the population, who are more financially sophisticated and invest their money in accordance with the CAPM alpha, whereas the remaining 83% of the population blindly follow unadjusted returns. Therefore, our micro-level results characterize the flow-performance relation of the median investor in the economy, whereas aggregated results focus on the relation of the representative agent. Moreover, we highlight the lack of financial sophistication among investors in DC markets, which has important policy implications.

To study how investors evaluate fund performance and make investment decisions, we hand-collected a large sample of DC 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 to study investors' asset allocation decisions when they face a bounded choice set so that the performance rankings of securities can be cleanly identified. Moreover, DC plans have grown in popularity and have become a large part of the market. Over 55 million American workers are active participants, with over $4.7 trillion 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, highlighting the importance of understanding the investment behavior of employees. The lack of evidence on the pattern and quality of investment decisions is disconcerting, given that employees rely heavily on their 401(k) to generate retirement income. As a result, our dataset provides an excellent resource for researchers and policymakers to understand investment decisions of the general public.

We begin the flow-performance analysis by estimating horse-race panel regressions of fund flow on various performance measures, such as the fund's net of fee returns (unadjusted returns), CAPM alpha, four-factor alpha, and Morningstar standardized returns. The regressions are similar to the horse-race panel regressions in Barber et al. (2016), but ours are estimated at the micro-level. Our triplet panel setting at the firm-fund-time level allows us to take into account the number of investment options that employees face when they allocate their savings. We show that investors' asset allocation decisions within the plan depend only on the unadjusted return. The result is in direct contrast to that of Barber et al. (2016), who suggest that investors are able to calculate the CAPM alpha and invest accordingly. Our finding demonstrates that even though the representative agent cares about the CAPM alpha, the median investor in the economy is less sophisticated and simply follows the funds'

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

past returns.

To assuredly establish the flow-return relation, we run two additional tests that are less parametric than the linear regression model, which alleviates the concern of the convex flow-performance relation that cannot be addressed in linear regression models. First, following Barber et al. (2016), we exploit the within-plan variations in performance rankings that result from different performance metrics used by investors. For example, a fund can be ranked in the top tier using unadjusted returns or in the middle tier using the CAPM alpha. We then run pairwise comparisons among alternative performance metrics. In the second non-parametric test, following Ben-David et al. (2019), we rank funds within the plan into quintiles based on each performance measure. For each performance metric, we examine the difference in flows between the top-quintile and bottom-quintile funds. Both methods show that flows are more responsive to funds with higher ranks of unadjusted returns than to those with higher ranks of CAPM alphas, four-factor alphas, or Morningstar ratings.

Investors blindly following unadjusted returns seems striking to economists, who are trained to take into account any systematic exposure when they evaluate performance. However, it is not necessarily the case that investors in our sample invest irrationally. It could simply be that following unadjusted returns maximizes investors' expected utility given their belief. To test this hypothesis, we assume employees in each plan as an agent with a mean-variance utility function, which depends on the agent's choice of performance measure. Given the assets in place from the prior year, the agent chooses the optimal asset allocation for the current year to maximize his expected utility. We then compare the distance between the actual asset allocation and the optimal one. We show that the actual asset allocation is closest to the optimal portfolio using unadjusted returns instead of the risk-adjusted alphas. Therefore, the observed asset allocation decisions are consistent with investors' belief that unadjusted returns serve as the ranking criterion.

Our results emphasize the flow-return relation at the micro-level, which is in stark contrast to the flow-alpha relation described in previous studies using fund-level data. One simple explanation for the contradicting evidence could be that the investors in our sample are drastically different from those in the existing studies. To rule out the sample selection bias, we aggregate our data at the fund level and re-run the horse race among alternative performance metrics. Interestingly, only the CAPM alpha can explain future flows at the aggregate level, which is consistent with findings in the existing literature. Therefore, our micro-level results are not driven by sample selections or a different clientele effect.

We show that the cross-sectional variation of wealth accumulation and investment strategy can explain why flows respond to the CAPM alpha at the aggregate level but respond to the risk-unadjusted return at the micro-level. We split our sample by 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 allocate their savings into funds based on the CAPM alpha, whereas employees in the remaining groups chase unadjusted fund returns. More importantly, the dispersion in average savings is enormous. Specifically, the highest-savings group represents only 17% of the population but holds more than 61% of the savings in our sample. The employees in the highest-savings group save on average $121,616, which is 27 times greater than those in the lowest-savings group and over 4.4 times greater than those in the second-highest-savings group. Therefore, the flow-performance relation at the micro-level is dominated by the majority of the population favoring unadjusted returns, but the relation tilts towards the CAPM alpha in the aggregate.

What are the driving forces of the savings gap, and why does the highest-savings group follow the CAPM alpha but the remaining groups follow unadjusted return? To answer these questions, we examine key determinants of savings in the 401(k) plan, including the matching contribution policy of the firm, work tenure of employees, plan performance, employees' wages, and their deferral rate.[2] First, the matching contribution policy of the firm allows employees to boost their savings with the additional contribution coming from employers' pockets. Second, employees with longer work tenures mechanically save more. Third, it could be that the highest-savings group has access to better investment plans than the lowest-savings group such that the wealth accumulates faster. We show that the first three determinants jointly explain only 25% of the savings gap, leaving a large proportion of the savings gap potentially explained by differences in wages and deferral rates.

We argue that the difference in financial literacy between the highest-savings group and the remaining population drives differences in wages and deferral rates, which, in turn, leads to a massive difference in overall savings. Such a gap in financial literacy also explains why wealthy employees are sophisticated enough to adjust market exposures when making investment decisions. 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)). In this paper, we use the State-by-State survey data from the FINRA Foundations National Financial Capability Study and show that employees with higher savings indeed reside in states with a higher financial literacy score. Specifically, a one standard deviation increase in the finance test score is associated with a 10% increase in employee savings. As a result, the highest-savings employees are more financially sophisticated, which explains why they direct their savings to funds with high CAPM alphas rather than unadjusted returns.

Lastly, we show that employees walk away from substantial capital gains by chasing un-

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

adjusted returns. For each plan at each year, we bootstrap hypothetical flows, assuming that investors direct flows based on the CAPM alpha rather than unadjusted returns. The observed flows underperform the hypothetical flows by 230% over 23 years, indicating that investors can benefit by taking into account market exposure when evaluating fund performance. In terms of economic magnitude, investors potentially leave \$1.3 billion on the table by overlooking the market exposure. However, both strategies significantly underperform the S&P 500 index fund available in most 401(k) plans. Our result suggests that unsophisticated investors are better off avoiding active funds and investing in index funds.

In summary, we find strong evidence that employees allocate their savings by following the unadjusted returns, which informs recent debates (Barber et al. (2016), Berk and van Binsbergen (2016), Ben-David et al. (2019), Evans and Sun (2018)). The literature have examined the flow-performance relation almost exclusively at the aggregate level, whereas our paper is the first to provide evidence at the micro-level. One advantage of our study is that we can identify the performance rankings faced by investors given their limited choices, which is impossible to achieve with the aggregated data. Unlike Ben-David et al. (2019) and Evans and Sun (2018), we explore the cross-section variation in investing patterns to explain the difference in results between our paper and that of Barber et al. (2016). We show that a small population in our sample holds a large proportion of wealth and favors the CAPM alpha over other performance measures, which drives the flow-performance relation towards the CAPM territory at the aggregate level.

Our paper also contributes to the wealth inequality literature and the rapidly growing financial literacy literature. The savings gap we identified can be explained by the difference in financial sophistication. Our finding is consistent with that of Lusardi et al. (2017), who show that financial knowledge is a crucial determinant of wealth inequality.

The rest of this study is organized as follows. Section 2 describes the dataset. Section 3 shows the flow-return relation at the micro-level. Section 4 discusses the difference in flow-performance relations between the micro-level and aggregate level. Section 5 concludes our discussion.

2 Data

2.1 Data sources

We have hand-collected employees' investments in 401(k) plans from the annual Form 11-K. This form is required by the U.S. Securities and Exchange Commission (SEC) if a firm offers company stock to employees in the plan. Our sample includes 1,551 public firms from 1993 to 2016. Firms report a list of securities available to their employees and the current investment value for each security. The list typically consists of the employer's stock, a set of mutual funds, and other securities, such as municipal bonds and stocks from other companies. For a few firms that provided multiple 401(k) plans for different subsidiaries in a given year, we aggregated them to one plan.

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 mutual funds in CRSP with the ones in the 11-K by fund names. Since most funds are listed on the 11-K without share class information, we aggregated the 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 the 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.[3] 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 firm-level. The average and median plan sizes are $370 million and $73 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 of all public and private firms that have plans with 100 or more participants.[4]

Employees invest a substantial proportion of capital in their firm's stock[5] in the early years of our sample. However, the share invested in company stock has declined gradually

[3]Fama-French 12 industries consist of business equipment, chemical, consumer durable, consumer non-durable, energy, finance, healthcare, manufacturing, telecommunication, utility, wholesale and retail, and others.

[4]Data on 401(k) assets for all firms are in Private Pension Plan Bulletin - Abstract of Form 5500 Annual Reports from the Department of Labor: https://www.dol.gov/agencies/ebsa/researchers/statistics/retirement-bulletins/private-pension-plan.

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

from 51% in 1993 to 24% in 2016, along with an increase in the number of available funds in the plan over the years. For example, an average of 17 funds are offered in 401(k) plans in 2015 compared to only three funds in 1993.[6] We then split funds into three categories: equity funds, bond funds, and a blend of these two fund types. For mutual fund investment, employees allocate 67% of their capital to equity funds and 20% to bond funds. These allocations are stable over the years examined.

3 Micro-level Flow Analysis

Even though both employees and employers contribute to 401(k) retirement plans, employees are responsible for making the investment decision. With the list of mutual funds provided by employers, employees allocate their capital across these options to generate retirement income. Therefore, our dataset provides a perfect context in which to study how retail investors evaluate fund performance and make investment decisions accordingly.

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

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

where 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, excluding stock holding.[7] Through out the paper, we winsorize flows at the 1% level

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

[7]Our results are robust if we also include stock holding in the denominator.

to remove outliers.

3.1 Panel regression

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

$$Flow_{pf,t+1} = \beta_0 PERF_{ft} + \mathbf{X'}_{pft}\boldsymbol{\beta}_1 + \mu_t + \gamma_p + \epsilon_{pf,t+1}, \tag{2}$$

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 (R_{ft}), [2] CAPM alpha (α_{ft}^{CAPM}), [3] 4-factor alpha ($\alpha_{ft}^{4Factor}$), and [4] Morningstar return (*MStar return*$_{ft}$) that is standardized within each investment category. $\mathbf{X}_{pft}$ represents plan and fund characteristics, which are the logarithm of the number of funds, the logarithm of the plan size, firm return, fund expense ratio, fund turnover ratio, the logarithm of total fund net assets, and standard deviation of fund return. μ_t and γ_p are time and firm fixed effects, respectively.

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 following time-series regressions using 36 months of returns:

$$R_{f\tau} - RF_\tau = a_{fm} + \mathbf{F}'_\tau \boldsymbol{\beta}_{fm} + \varepsilon_{f\tau}, \quad \tau = m-1, ..., m-36 \tag{3}$$

where RF_τ is the risk-free rate in month τ 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 which 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 et al. (2019), Cici and Gibson (2012), and Elton et al. (1995). We then estimate alpha for fund f in month m as follows:

$$\hat{a}_{fm} = R_{fm} - RF_m - \mathbf{F'_m}\hat{\boldsymbol{\beta}}_{fm}, \tag{4}$$

where $\hat{\boldsymbol{\beta}}_{fm}$ is estimated from equation (3). Since our data are at an annual frequency, an annual alpha of fund f is calculated using the 12 monthly alpha in year t:

$$\alpha_{ft} = \prod_{j=0}^{11}\left(1 - \hat{a}_{f,t-\frac{j}{12}}\right) - 1. \tag{5}$$

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[8] f at time t as follows:

$$MRAR_{ft}(\gamma, T) = \left[\frac{1}{T}\sum_{j=0}^{T-1}(1 + ER_{f,t-j})^{-\gamma}\right]^{-\frac{12}{\gamma}} - 1, \tag{6}$$

where γ 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

[8]The Morningstar Rating for Funds is available at https://s21.q4cdn.com/198919461/files/doc_downloads/othe_disclosure_materials/MorningstarRatingforFunds.pdf.

in a given investment category (i.e., the top 10% will be 5-star funds, and the bottom 10% will be 1-star funds).

Columns (1) to (4) in Table 2 demonstrate that each of these four performance measures significantly predicts future flows. To compare predictive powers across models, we first make pairwise comparisons between the fund's net return and other measures. Column (5) reports the regression results of future flows on net return and CAPM alpha. It shows that the coefficient on the CAPM alpha not only becomes much smaller, but it is also insignificant, while the coefficient on net return is 0.016 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). In addition, coefficients on net return stay almost the same across pairwise comparisons with an average value of 0.02. Our finding suggests that when fund net return increases by 1%, there is additional 0.02% of flow, which is equivalent to $40,000 extra for funds in the average firm's plan.

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. In column (9), we add the investment category fixed effects, and our results are robust.

3.2 Pairwise model horse race

In the previous section, we examined the flow-performance sensitivity assuming a linear relation. To relax the linear relation assumption and follow 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:

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

where D_{ijpft} is a dummy variable that equals one if fund f of firm p in year t is in quantile i based on the net return measure and quantile j based on alternative performance measure, such as the CAPM alpha or the Morningstar standardized return. $\mathbf{X}_{pft}$ represents a vector of control variables used in equation (2). Firm and year fixed effects are also included.

To estimate the model, we exclude the dummy variable for $i = 3$ and $j = 3$. By excluding the dummy variable, b_{ij} represents the percentage flows to a fund in quantile i and j based on fund net return and other return measures, respectively, relative to a fund that was in the third quantile for both performance measures. Investors use net returns to direct their flows if the sum of differences in coefficients on dummy variables $(b_{i,j} - b_{j,i})$ for all i and j such that $i > j$:

$$\text{Sum} = \sum_{i>j}^{5} \sum_{j=1}^{4} (b_{i,j} - b_{j,i}) \tag{8}$$

is significantly greater than zero.

Table 3 reports the sum of differences in coefficients for the horse races between the return model and the others. The results show that investment flows to funds that have higher unadjusted returns instead of funds with a higher CAPM alpha, 4-factor alpha, or Morningstar standardized returns. Moreover, for each horse race, we calculate the fraction of positive coefficient differences $(b_{i,j} - b_{j,i})$ and conduct the hypothesis test that the fraction equals 50%. The results show that these coefficients are positive 90-100% of the time. Therefore, our pairwise comparison results are not driven by some extreme estimates of coefficients.

3.3 Top-ranked and bottom-ranked funds

This section provides evidence that investors follow returns, using a non-parametric approach developed in Ben-David et al. (2019). Within each plan, we rank funds into quintiles by unadjusted returns, the CAPM alpha, 4-factor alpha, and Morningstar standardized returns,[9] respectively. We then compare flow differences between the top-quintile and the bottom-quintile 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 difference in percent flows between the top and bottom return quintiles is 0.72%, which is statistically significantly higher than 0.41% for Morningstar quintiles, or 0.46% for the CAPM alpha quintiles. In addition, the difference in the fraction of positive flow between the top and bottom return quintiles is 6.31%, which is significantly higher than that of other performance quintiles. Finally, the ranking using net return generates the highest spread in dollar flows of $640,378 compared to other rankings with the average spread of $217,083.

3.4 Mean-variance portfolio

We have shown that investors follow raw returns when making investment decisions. Chasing unadjusted-returns is not irrational if investors maximize their expected utility, given their belief that net return is the relevant measure to evaluate fund performance. To test this

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

hypothesis, we assume that employees in each plan act as a representative agent with a mean-variance utility function. Given his or her assets in place from the prior year, the agent chooses the best asset allocation at each year to maximize the utility.

Formally, employees at firm p in year t allocate their capital to each fund f with proportion of wealth $\mathbf{w}'_{pt} = [w_{f_1}, w_{f_2}, ..., w_{f_{N_{pt}}}]_t$, where N_{pt} is the number of funds that are in firm p's 401(k) plan. Their optimal allocations for next period $t+1$ are the solutions of:

$$\max_{\mathbf{w}_{p,t+1}} \frac{\mathbf{w}'_{p,t+1} E_t[R^e_{p,t+1}]}{\mathbf{w}'_{p,t+1} \boldsymbol{\Sigma}_{p,t+1|t} \mathbf{w}_{p,t+1}} \tag{9}$$

$$\text{st: } \mathbf{1}'\mathbf{w}_{p,t+1} = 1, \tag{10}$$

$$\mathbf{w}_{p,t+1} \leq (1+b)\mathbf{w}_{pt}, \tag{11}$$

$$\mathbf{w}_{p,t+1} \geq \max\{(1-b)\mathbf{w}_{pt}, 0\}, \tag{12}$$

where $\mathbf{1}$ is vector of 1, and b is the boundary constraint obtained from the historical cross sectional and time series distribution of changes in allocations.

The expected returns in excess of the risk-free rate $E_t[R^e_{p,t+1}]$ and the conditional covariance of returns $\boldsymbol{\Sigma}_{p,t+1|t}$ are estimated under different investor's belief or utilization (i.e., unadjusted return, the CAPM alpha, or 4-factor alpha). Under this belief, the excess returns will follow either:

- Net return: $$R^e_{f,t+1} = \frac{1}{36} \sum_{j=0}^{35} R^e_{f,t-j} + \varepsilon_{f,t+1}, \tag{13}$$

- CAPM or 4Factor: $$R^e_{f,t+1} = a_{ft} + \mathbf{F}'_{\mathbf{t+1}} \boldsymbol{\beta}_{ft} + \varepsilon_{f,t+1}, \tag{14}$$

where $\varepsilon_{f,t+1}$ has a conditional normal distribution that has mean zero and conditional variance $\sigma^2_{f,t+1}$, or $\varepsilon_{f,t+1}|_{I_t} \sim N(0, \sigma^2_{f,t+1})$. The variance $\sigma^2_{f,t+1}$ follows GARCH(1,1) (Bollerslev (1986)):

$$\sigma^2_{f,t+1} = w_{1f} + w_{2f}\varepsilon^2_{ft} + w_{3f}\sigma^2_{ft}. \tag{15}$$

The conditional covariance of returns of all funds in plan p in year t are defined as follows:

$$\mathbf{\Sigma}_{p,t+1|t} = \mathbf{D}_{p,t+1|t}\mathbf{Q}_{p,t+1|t}\mathbf{D}_{p,t+1|t}, \tag{16}$$

where $\mathbf{Q}_{p,t+1|t}$ is the correlation matrix estimated using past 36 months of fund return data, and $\mathbf{D}_{p,t+1|t}$ is the diagonal matrix with the standard deviation $\sigma_{f,t+1}$ along the diagonal.

Under each model, for all funds f in firm p's 401(k) plan at year t we estimate conditional expected excess returns and covariance matrix using the past 36 months of return data. Thus, the expected excess returns are:

- Net return: $$E_t[R^e_{f,t+1}] = \frac{1}{36}\sum_{j=0}^{35} R^e_{f,t-j}, \tag{17}$$

- CAPM and 4Factor: $$E_t[R^e_{f,t+1}] = E_t[\mathbf{F}'_{t+1}]\hat{\boldsymbol{\beta}}_{f,t}, \tag{18}$$

where the expected factor returns are $E_t[\mathbf{F}_{t+1}] = \frac{1}{36}\sum_{j=0}^{35}\mathbf{F}_{t-j}$, and the conditional covariance matrix is $\mathbf{\Sigma}_{p,t+1|t} = \hat{\mathbf{D}}_{p,t+1|t}\hat{\mathbf{Q}}_{p,t+1|t}\hat{\mathbf{D}}_{p,t+1|t}$.

Given the estimated expected returns and covariance of returns along with prior allocations from last year, the employees maximize their utility in equation (9) to obtain the optimal portfolio. We then compare the distance between the actual asset allocations and the optimal holdings using these two metrics: [1] Δ_{pt} is the average of the absolute difference between the optimal allocations (w^{Model}_{pkt}) and the actual ones (w^{Actual}_{pkt}), and [2] D_{pt} is the distance between them :

$$\Delta_{pt}^{model} = \frac{1}{N_{pt}} \sum_{k=f_1}^{f_{N_{pt}}} \left| w_{pkt}^{Model} - w_{pkt}^{Actual} \right|, \tag{19}$$

$$D_{pt}^{model} = \sqrt{\sum_{k=f_1}^{f_{N_{pt}}} \left(w_{pkt}^{Model} - w_{pkt}^{Actual} \right)^2}. \tag{20}$$

To estimate optimal allocations, we use three different boundary constraints (b): 8%, 20%, and 40%. These constraints are obtained from historical observations of the absolute changes in allocations over time. These numbers are at the 25^{th}, 50^{th}, and 75^{th} percentile of the distribution, respectively. Under each boundary constraint, Panel 1 in Table 5 shows that the averages of the absolute difference between the observed holdings and the optimal holdings estimated using the net return model range from 2.82% to 5.36%. The difference is significantly smaller than those obtained using other models. Panel 2 provides the results for the distance between the vector of actual allocations and that of optimal ones. Regardless of boundary constraints, optimal allocations using net return are significantly closer to the actual holdings when compared to the optimal allocations obtained using other models. Therefore, the result suggests that the observed asset allocation decisions are consistent with investors belief that unadjusted returns serve as the ranking criteria.

4 Aggregated Flow Analysis

To reconcile the difference between our micro-level results and the results in the literature (Barber et al. (2016), Berk and van Binsbergen (2016)), we aggregate our data at 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

$$AGGflow_{ft} = \frac{V_{ft} - V_{f,t-1}(1 + R_{ft})}{V_{f,t-1}}, \tag{21}$$

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:

$$AGGflow_{f,t+1} = \delta_0 PERF_{ft} + \mathbf{Z'}_{ft}\boldsymbol{\delta} + \mu_t + \eta_{f,t+1}, \tag{22}$$

where $\mathbf{Z'}_{ft}$ represents fund characteristics which are the fund expense ratio, fund turn over, logarithm of total fund net assets, and standard deviation of fund returns.

The results are shown in Table 6. The first four columns show that aggregated flows respond to each of the four performance metrics, except for the Morningstar return. Column (5) runs a horse-race among all performance measures, and only the CAPM alpha is positive and significant, which is consistent with the prior literature (Barber et al. (2016), Berk and van Binsbergen (2016)). Our results also alleviate the concern that the difference between our micro-level results and findings in the literature is driven by sample selection.

4.1 Reconciling the aggregated and micro-level analyses

Why do aggregated flows respond to the CAPM alpha, while the micro-level flows respond to the unadjusted returns? We conjecture that the inconsistency is driven by different weighting schemes. In particular, the micro-level evidence suggests that the "average" employee in the firm allocates savings by following unadjusted returns, and the analysis effectively puts more weight on firms with a small plan size. In contrast, the aggregated analysis puts more weight on firms with a large plan size. To reconcile the differences in results between our research and prior studies, we offer a potential explanation using the wealth inequality channel. 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, we rank firms into quartiles by the average savings per employee of the plan. For each quartile, we re-run the micro-level panel regression of flows on performance measures. The results are shown in Table 7. Almost all quartiles show a flow-return relation, except for the highest-savings quartile, where flows only respond to the CAPM alpha. Furthermore, by examining the plan characteristics of these subsamples, Table 8 shows that the average plan size of the highest-savings group is $840 million, while that of the other groups has a significantly smaller value of $201 million. The highest-savings group also has significantly fewer employees. In fact, this group represents only 17% of the population in the sample, but holds more than 61% of the savings. Therefore, these results confirm our hypothesis.

4.2 Contribution channels to difference in savings

Table 8 shows that employees of the highest-savings group save an average of $121,616 per capita, which is over 7.8 times greater than the average savings of other groups. What factors contribute to this difference in savings? To answer these questions, we define the total savings in the 401(k) plan of each employee up to time t as follows (see the Appendix for the derivation):

$$Saving_{1\to t} = Salary_t \frac{(1+g)^t - (1+r)^t}{(g-r)(1+r)^{t-1}} \left[s + min\{f_{cap}, s \times f_{rate}\}\right] \tag{23}$$

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 (23), there are five factors that could potentially contribute to the savings gap between the highest-savings group and all other 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 firm's matching policy only explains 1% of the savings gap. Table 8 shows that firms will match up to 5.48% of employee compensation for the highest-savings group and 5.35% for the remaining groups. The difference in contribution cap is statistically significant but economically insignificant. In terms of contribution ratio, employees from the highest-savings group receive from their employers an extra $0.73 for every dollar they contribute to the plan compared to $0.62 for employees from the remaining groups. Even though the difference of $0.11 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 highest-savings group will have 7% more capital compared to those in the other groups. The 7% additional capital accounts for only 1% of the difference in savings of $106,024 between the highest-savings group and the remaining groups.

Second, the massive savings gap in the 401(k) cannot be entirely attributed to the long working tenures of the highest-savings group. We use the difference in firm ages as a proxy for the difference in working tenures between the highest-savings employees and the others. Firm age is defined as the time since we have firm data from either the CRSP or Compustat datasets. The average difference in firm ages 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 $101,202.[10] However, the income per capita in the U.S. is $46,550 in 2016,[11] suggesting that the savings gap is unlikely to be due to the difference in work tenure. Furthermore, the 10-year gap in firm ages is likely to overestimate the actual tenure gap, making the tenure gap even more unlikely to explain the savings gap.[12] In the Appendix, 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 (23) 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 8. The plan size growth rates are set to the average

[10]Using equation (23) 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.

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

[12]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 age from 55 to 64 is 10.3, whereas that of younger workers with age 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 highest-savings group is full of workers with age between 55 and 64, and the firm in the lowest savings group is full of workers with age 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.

performance of plans from 1993 to 2016, which are 6.4% and 5.2% for the highest-savings group and the others, 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 19% of the difference in savings, and the plan's investment return difference contributes another 5%.

Therefore, the difference in savings is mainly driven by the difference in employees' deferral rate and their annual compensations. To determine the contributions of these two effects, we estimate the ratios of the deferral rate s and the $Salary_t$ between the highest-savings group and the others using equation (23) with the tenure gap of 9.4 years. Figure 1 illustrates the combination of these two ratios that help explain the remaining 75% of the savings gap. For example, if employees from these two groups have the same salary and those who are not from the highest-savings group saved 3% per year,[13] employees in the highest-savings group 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 employees in the highest-savings group would earn annually around \$139,650, assuming employees in the other group earned the average income per capita 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.

4.3 Financial literacy affects savings

We argue that financial literacy is strongly related to the deferral rate and compensation. Hence, it is a key driver of the difference in savings. Furthermore, such a gap in financial literacy also explains why the highest-savings employees are sophisticated enough to adjust market exposures when making investment decisions.

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

To formally link deferral rate and compensation to financial literacy, we use the State-by-State survey data from the FINRA Foundation's National Financial Capability Study. 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. Figure 2 (a) and (b) show that respondents with higher test scores are more likely to make regular contributions to their retirement plans. Furthermore, the higher the recurrence of contributions, the higher the average 401(k) deferral rate. Consequently, deferral rates correlate with the level of financial sophistication. In addition, Figure 2 (c) and (d) show that compensation is strongly associated with financial literacy.

Figure 2 provides descriptive evidence of the positive correlations among the financial literacy, the deferral rate, and the compensation. To formally test whether employees with high savings are also financially sophisticated, we run regressions of the logarithm of average 401(k) savings per employee of firm p that is located in ZIP code z of state s on the six determinants from equation (23) and firm characteristics as follows:

$$\begin{aligned}
\ln(Saving_{pt}) = {} & \beta_1 Firm\ matching\ ratio_{pt} + \beta_2 Firm\ contribution\ cap_{pt} \\
& + \beta_3 Firm\ age_{pt} + \beta_4 Plan\ return_{pt} + \beta_5 Employee\ salary_{zt} \\
& + \beta_6 Employer\ contribution\ rate_{zt} + \beta_7 Finance\ test\ score_{st} \\
& + \mathbf{M'}_{pt}\boldsymbol{\beta}_8 + \mu_t + \xi_{pt}, \qquad (24)
\end{aligned}$$

where $Firm\ age_{pt}$ is the logarithm of firm age; $Plan\ return_{pt}$ is the value-weighted 401(k) plan return. $\mathbf{M'}_{pt}$ represents firm characteristics which are the total debt to total asset, the book equity value to market equity value, the logarithm of market equity value, and the gross margin to a total asset.

From equation (23), *Firm matching ratio* and *Firm contribution cap* have non-linear relationships with *Saving*. In particular, *Saving* increases linearly with *Firm matching ratio* up to the value of $\frac{\textit{Firm contribution cap}}{\textit{Employee deferral rate}}$, and remains flat once *Firm matching ratio* is over the threshold. A similar pattern also holds for *Firm contribution cap*. Therefore, to capture the non-linear relations for both *Firm matching ratio* and *Firm contribution cap* with *Saving*, we transform the two variables using a transformation function $f(x) = -e^{-x}$.

To proxy for employee salary ($\textit{Employee salary}_{pt}$) at firm p that is located in ZIP code z, we use the average individual salary ($\textit{Employee salary}_{zt}$) of those who live within 25 miles away of firm p's headquarter. Specifically, data on personal wages at the county level are obtained from the Internal Revenue Service (IRS). We then convert the county code into the ZIP code. The $\textit{Employee salary}_{zt}$ is the logarithm of the average individual salary of those who live in ZIP codes where the great-circle distances between those codes and ZIP code z are within 25 miles.[14] To proxy for employee deferral rate at firm p that is located in ZIP code z, we use the average contribution to pension rate ($\textit{Employer contribution rate}_{zt}$) of all employers who establish in ZIP code z. The contribution to pension rate is the ratio of the employer's contributions in an employee pension to individual wages. The county-level data for these variables are obtained from the Bureau of Economic Analysis from 1969.[15] Finally, the state-level data for the $\textit{Finance test score}_{st}$ are the percentage of corrections on the financial literacy test from FINRA.

Table 9 shows the effects of the determinants on employee savings. The six main determi-

[14] County-ZIP conversion data is from HUD USER. ZIP codes distance data are from Gazetteer and U.S. Census SF1.

[15] We use the similar procedure of getting the employee salary at ZIP code z to get employer contribution rate at ZIP code z.

nants of employee savings have expected signs, and most of them are statistically significant. The firm's matching policy strongly correlates with employee savings. For median firms, a one standard deviation increase in either the firm's matching ratio or the firm's contribution cap is associated with an 15% or 10% increase in savings, respectively.[16] Both firm age and plan return have strong positive relations with the savings as they should be.

Employees of firms located in areas with high salaries have higher savings at ten-percent significance level. For example, for every 10% increase in salary, employees will have 4.9% more in their savings. In addition, using the employer's contribution rate to proxy for the employee's deferral rate, the results show that the employer's contribution rate does not drive employee savings. It is possible that the employer's contribution rate has a low correlation with the employee's deferral rate in the early years of our sample. This is because the employer's pension contribution accounts for both defined benefit and defined contribution plans. Furthermore, firms were more likely to provide defined benefit plans prior to 2000, and then shift to defined contribution plans afterwards.

Finally, the results show that employees who reside in the states where the residents perform better on the finance test have significantly higher savings. Specifically, a one standard deviation increase in a finance test score is associated with a 10% increase in employee savings. Consequently, the highest-savings group has significantly better financial capability than the remaining groups, which explains that the highest-savings group selects funds based on the CAPM alpha.

[16]The total effects of firm contribution policy on savings are economically stronger than what we discuss in section 4.2. It is because the difference in firm contribution policies between the highest saving group and the others is much smaller than one standard deviation change in values. If we consider the difference values between these groups, the results from Table 9 show that the firm matching ratio and the firm contribution cap contribute to 6% and 1% difference in savings.

4.4 Welfare lost

Having documented that investors in the low-savings/low-financial-capability groups follow unadjusted returns, we show that these investors give up substantial capital gains without adjusting for market exposure. We consider the following two counterfactuals. First, for each firm-year panel, we bootstrap hypothetical flows, assuming that investors direct flows based on the CAPM alpha rather than unadjusted returns. Second, we assume that investors simply invest passively in S&P 500. At the end of each year, we compute the performance of money flows to 401(k) plans over the next year:

$$R_{t+1} = \sum_{f \in \Omega_t} w_{ft} R_{f,t+1}, \tag{25}$$

$$w_{ft} = \frac{\sum_{p \in \Omega_{ft}} \$Flow_{pft}}{\sum_{f \in \Omega_t} \sum_{p \in \Omega_{ft}} \$Flow_{pft}}, \tag{26}$$

where Ω_t is the set of funds f that are available in all 401(k) plans in year t and Ω_{ft} is the set of plans p that have fund f in their menu in year t. The cumulative performance of flows is the compounded performance of flows over the years. We then compare cumulative performances of the observed flows, the hypothetical flows, and the S&P 500 index.

To simulate the hypothetical flows, we assume that investors chase the CAPM alpha as follows:

$$Flow_{pft} = \beta_0 \alpha^{CAPM}_{f,t-1} + \mathbf{X'}_{pf,t-1} \boldsymbol{\beta}_1 + \mu_t + \gamma_p + \epsilon_{pft}, \tag{27}$$

We then run the residual-resampling bootstrap of equation (27) to simulate the percentage flow $Flow^s_{pft}$ for each fund-plan. Next, the percentage flow is converted to the dollar flow $\$Flow^s_{pft}$.[17] We then aggregate the hypothetical flows and compute their cumulative

[17] Within each plan, we adjust the simulated dollar flow $\$Flow^s_{pft}$ by a same factor so that the sum of

performances over the sample period. We repeat these procedures 1,000 times.

Figure 3 shows the cumulative performances of the observed flows, the hypothetical flows, and the S&P 500 index fund over the sample period. The observed flows significantly underperform the hypothetical flows over our sample periods. Specifically, the observed flows earn around 227%, whereas the hypothetical flows earn 456% between 1994 and 2017. The difference in dollar terms is about $1.3 billion over the sample periods. The result suggests that investors can benefit from taking into account the market exposure when evaluating fund performance. Furthermore, both observed and hypothetical flows underperform the S&P 500 index fund, which is commonly available in DC plans. Our results suggest that unsophisticated investors are better off avoiding active investing and simply investing in market index funds.

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 micro-level, where investment options are naturally bounded. Using our hand-collected dataset on a large sample of defined contribution plans, we show that the majority of investors blindly follow unadjusted returns and these investors have substantial capital loss for not adjusting market exposure when making investment decisions. The result is disconcerting given that a large portion of pension assets are in 401(k) plans.

simulated dollar flows is equal to the sum of observed dollar flows.

Our paper has important implications to plan sponsors and investors. From the sponsors' perspective, it is necessary to offer a diversified set of funds to their employees. More importantly, sponsors can hold educational workshops, so that employees can better understand trade-offs between risk and return. It will also improve the stability of the plan when the market becomes volatile. From the employees' perspective, they are better off holding a well-diversified index fund rather than chasing high past returns.

References

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.

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.

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.

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

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, *Available at SSRN 3287997* .

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.

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

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.

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

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.

Appendix

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:

$$
\begin{aligned}
&\text{1. Plan growth:} && PG_2 = Saving_{1\to 1} \times G \\
& && \quad = Salary_1 \times Saving\ rate \times G, && (28) \\
&\text{2. New contribution:} && NC_2 = Salary_2 \times Saving\ rate \\
& && \quad = Salary_1 \times R \times Saving\ rate, && (29) \\
&\Rightarrow \text{Total saving:} && Saving_{1\to 2} = Salary_1 \times Saving\ rate \times \big(G + R\big) \\
& && \quad = Salary_1 \times Saving\ rate \times \frac{G^2 - R^2}{G - R}, && (30)
\end{aligned}
$$

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

$$Saving_{1\to t} = Salary_1 \times Saving\ rate \times \frac{G^t - R^t}{G - R}. \tag{31}$$

QED.

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 working tenures. Our inferred age difference between the highest-savings group and the remaining 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:

$$Age^{TA}_{pft} = 65 - (T_f - t) \tag{32}$$

Panel 1 in Table A1 shows that the difference in employee age between the highest and lowest savings groups is 1.6 years 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:

$$Age^{O}_{pft} = \Phi\Big(\mathbf{N'}_{pft}\boldsymbol{\beta} + \mu_t + \gamma_p\Big) + \zeta_{pft}, \tag{33}$$

$$Age^{O}_{pft} = \frac{Age^{K}_{pft} - 22}{65 - 22} \in (0, 1), \tag{34}$$

and $K = \{TA, NT\}$; $\Phi(.)$ is the cumulative distribution function of the standard normal distribution, and ζ_{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, turn-over ratio, the fund's return volatility, and the fund's time-varying loading on market, size, and value factors. Panel 2 in Table A1 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 highest group and the remaining groups is 1.6 years, which is clearly not a primary contribution to the difference in savings.

Figure 1
Contributions of salary and deferral rate to the savings gap

The figure shows the combination of deferral rate ratio and salary ratio between the highest-savings employees and the others. The deferral rate is the annual savings in a 401(k) plan as a percentage of annual salary. These combinations explain the difference in total savings among these employees in addition to the differences in the firm's matching contribution policy, the employee tenure, and the plan's investment returns.

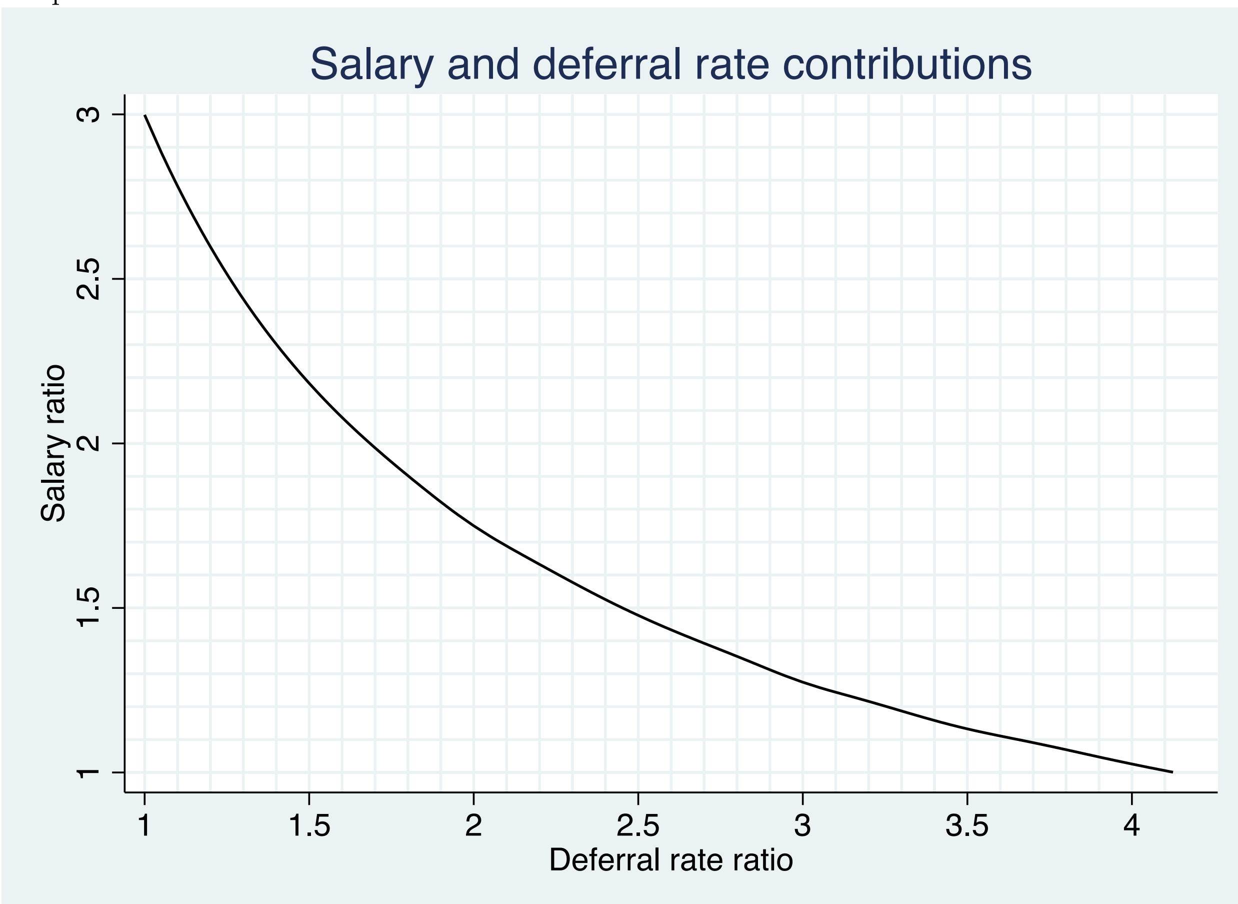
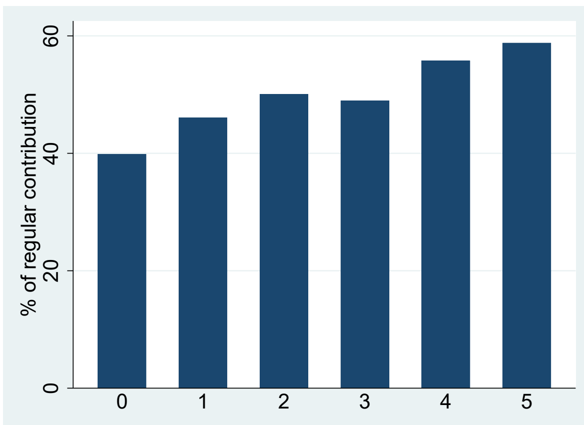
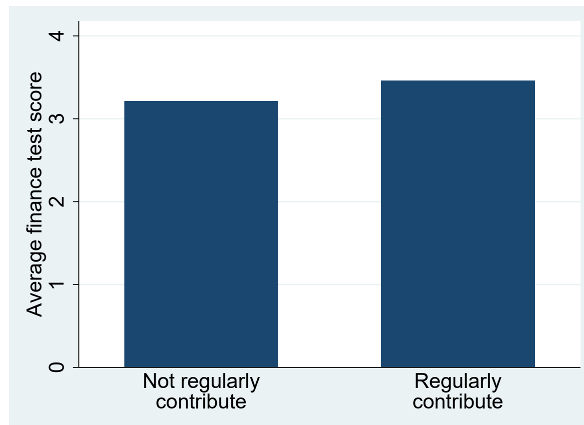


Figure 2
Financial literacy

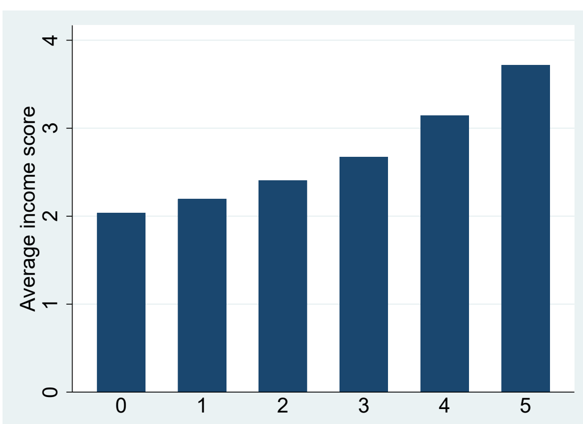
Figure (a) and (b) show the relation between the finance test score and the regularity of the contribution to the retirement plans. Figure (c) and (d) exhibit the relation between the finance test score and the income score. The income score ranges from 1, for those who have income less than $25,000, to 6, for those who earn at least $150,000.



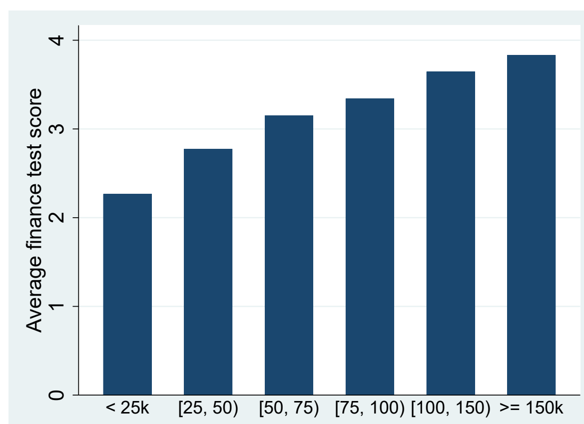
(a) Percentage of regular contribution to retirement plans by finance test score.



(b) Average finance test score by contribution levels.



(c) Average income score by finance test score.



(d) Average finance test score by income level.

Figure 3
Cumulative performance of flows to 401(k) plans

The figure shows the cumulative performance of flows to 401(k) plans. The black dashed line represents the cumulative returns of observed flows to 401(k) plans in our sample. Specifically, at the end of each year, we calculate the performance of flows over the next year. We also bootstrap hypothetical flows to provide a counterfactual scenario, in which investors follow the CAPM alpha ranking. The red solid line shows the average cumulative returns of the hypothetical flows to 401(k) plans over 1,000 simulations. The dotted line shows the 95% confidence intervals of the hypothetical flows. The blue long dashed line plots the cumulative returns of S&P 500.

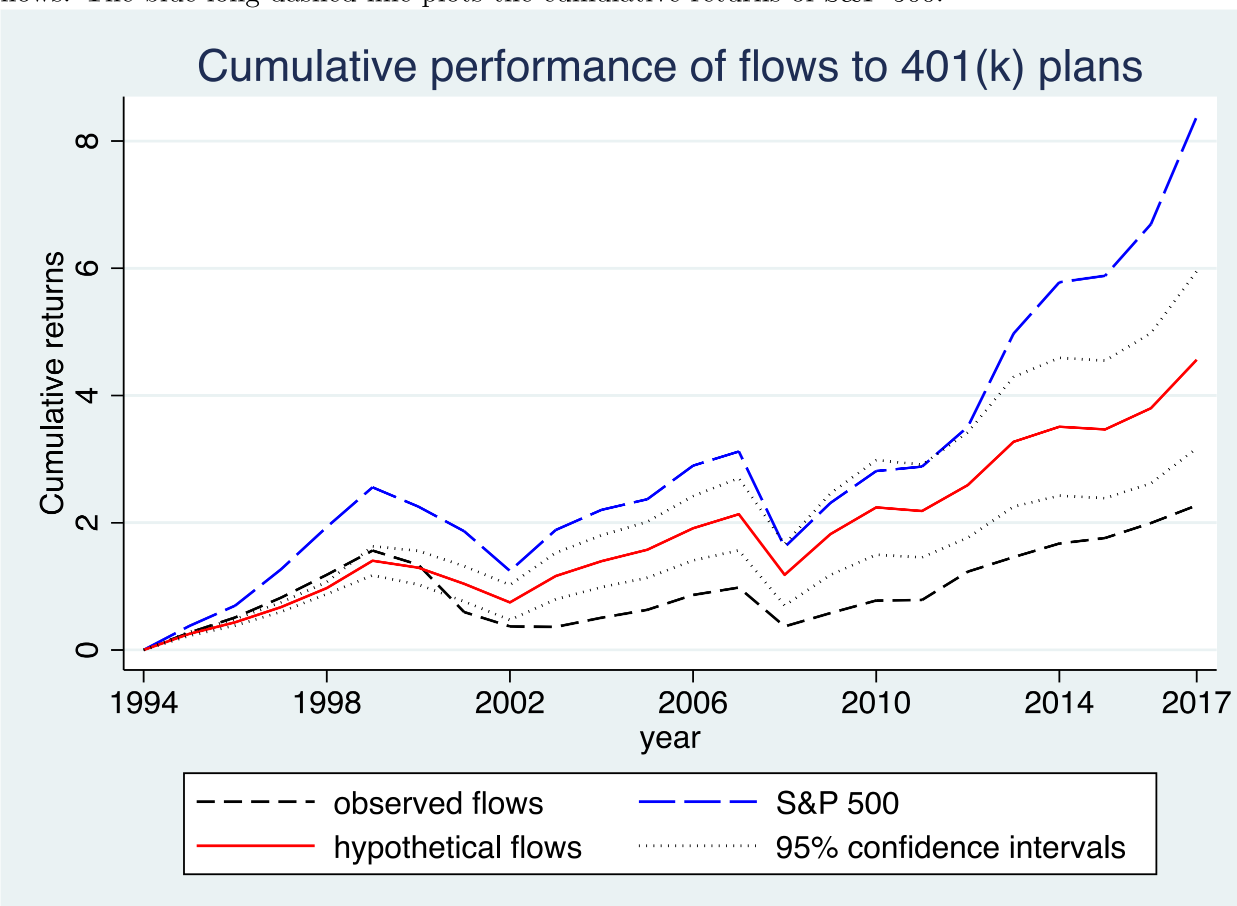


Table 1
Summary statistics

This table reports the descriptive statistics for our sample. Column 1 reports the number of firms with 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 shows the total number of mutual funds that are invested in all plans and the average number of funds used in each plan, respectively. Columns 7 and 8 report 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	90	416	36	50.93	110	3	60.24	22.71
1994	160	270	36	47.57	167	3	58.90	27.49
1995	244	178	27	47.12	242	3	58.70	27.67
1996	317	192	26	46.81	325	4	61.88	26.10
1997	404	243	29	41.20	435	4	64.76	23.46
1998	476	240	31	35.79	550	5	66.68	21.91
1999	528	237	42	31.45	722	6	71.24	18.41
2000	601	319	38	31.26	858	7	72.87	17.39
2001	683	250	37	32.84	1,022	8	67.73	21.09
2002	740	198	35	32.54	1,142	9	62.70	25.36
2003	812	267	47	34.09	1,304	10	67.19	21.93
2004	834	285	59	32.99	1,390	10	70.19	18.68
2005	810	372	69	31.20	1,438	11	71.87	17.58
2006	779	439	80	30.94	1,427	11	72.44	17.14
2007	768	479	84	27.21	1,447	12	71.45	16.77
2008	740	290	59	25.61	1,450	13	62.27	24.31
2009	724	352	74	25.25	1,424	13	63.81	22.48
2010	682	438	95	24.90	1,425	14	65.04	20.56
2011	668	445	99	24.00	1,437	15	62.43	21.46
2012	640	490	126	23.61	1,426	15	62.14	21.05
2013	611	631	151	24.67	1,612	16	67.14	16.46
2014	586	647	160	24.10	1,671	17	67.78	15.21
2015	542	587	157	21.49	1,621	17	68.05	14.71
2016	475	628	166	24.40	1,163	14	68.02	14.51

Table 2
Flow-performance relation at the firm-fund level

This table reports results from regressions of flows at the firm-fund level in year $t+1$ on fund performance, plan and fund characteristics in year t. There are four different measures of fund performance: [1] $Net\ return_t$ is fund return net of fee; [2] α_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. Plan characteristics include logarithm of the number of funds, firm return, and firm stock flow to a 401(k) plan. Fund characteristics include expense ratio, turn over, the logarithm of total fund net assets, and fund return volatility. Standard errors are two-way clustered at the firm and year levels. The t-statistics are in brackets. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
	$Flow_{t+1}$	$Flow_{t+1}$	$Flow_{t+1}$	$Flow_{t+1}$	$Flow_{t+1}$	$Flow_{t+1}$	$Flow_{t+1}$	$Flow_{t+1}$	$Flow_{t+1}$
$Net\ return_t$	0.019***				0.016**	0.020**	0.018**	0.017**	0.019**
	[2.87]				[2.54]	[2.65]	[2.74]	[2.65]	[2.69]
α_t^{CAPM}		0.022**			0.007			0.014	0.013
		[2.51]			[0.95]			[1.67]	[1.60]
$\alpha_t^{4Factor}$			0.014**			-0.004		-0.014	-0.015*
			[2.10]			[-0.55]		[-1.60]	[-1.77]
$MStar\ return_t$ ($\times 0.1$)				0.014**			0.004	0.003	0.004
				[2.37]			[0.81]	[0.54]	[0.72]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Invt. type FE	No	No	No	No	No	No	No	No	Yes
Observations	119,307	119,307	119,307	119,307	119,307	119,307	119,307	119,307	119,307
Adjusted R^2	0.17	0.16	0.16	0.16	0.17	0.17	0.17	0.17	0.17

Table 3
Pairwise horse-race

This table reports the pairwise horse-race between the net return model and the other competing models. Section 3.2 describes the estimation of following regression:

$$Flow_{pf,t+1} = \sum_i \sum_j b_{ij} D_{ijpft} + \mathbf{X'}_{pft}\mathbf{c} + \mu_t + \gamma_p + \epsilon_{pf,t+1},$$

where D_{ijpft} is a dummy variable that equals one if fund f of firm p in year t is in quantile i based on the fund net return and quantile j based on other performance measures. To estimate the model, we exclude the dummy variable for $i = 3$ and $j = 3$. The matrix $\mathbf{X}_{pft}$ represents firm and fund control variables, which are the same set of controls in Table 2. μ_t and γ_p are year and firm fixed effects. This table reports the sum of the differences between coefficients b_{ij} and b_{ji} for all i and j such that $i > j$:

$$\text{Sum} = \sum_{i>j}^{5} \sum_{j=1}^{4} (b_{i,j} - b_{j,i}).$$

The t-statistics are in brackets. ** and *** indicate significance at the 5% and 1% levels, respectively.

Winning Model	Net return	Net return	Net return
Losing Model	CAPM	4Factor	Morningstar
Sum of coefficient differences	0.019***	0.024***	0.013***
	[6.34]	[9.64]	[6.09]
% of coefficient differences > 0	100***	100***	90**
Binomial p-value	(0.002)	(0.002)	(0.021)

Winning Model	CAPM	Morningstar
Losing Model	4Factor	4Factor
Sum of coefficient differences	0.012***	0.007***
	[3.92]	[3.13]
% of coefficient differences > 0	90**	80
Binomial p-value	(0.021)	(0.109)

Tie Model	CAPM
Tie Model	Morningstar
Sum of coefficient differences	0.003
	[0.16]
% of coefficient differences > 0	40
Binomial p-value	(0.754)

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 quintiles. The top quantile and bottom quantile contain the best and the worst performing funds. *Positive Flows* are 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" quantiles and other quantiles. The t-statistics are in brackets. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Positive Flows (%)			Flows (%)			Dollar Flows (thousands)		
	Top	Bottom	Diff	Top	Bottom	Diff	Top	Bottom	Diff
Net return	67.00	60.69	6.31	1.38	0.67	0.72	856.63	216.26	640.38
CAPM	65.55	60.48	5.07	1.20	0.74	0.46	642.61	317.76	324.84
4Factor	64.81	62.35	2.46	1.12	0.78	0.35	568.62	371.88	196.75
Morningstar	66.51	61.57	4.94	1.16	0.76	0.41	539.30	409.64	129.66
The difference in *Diff* between Net return model and other models:									
- CAPM			1.24**			0.26***			315.54***
			[2.48]			[5.26]			[5.34]
- 4Factor			3.85***			0.37***			443.63***
			[6.98]			[7.41]			[6.84]
- Morningstar			1.37**			0.31***			510.72***
			[2.40]			[5.82]			[7.97]

Table 5
Distance to mean-variance portfolios

This table reports the average of the absolute difference (Δ_{pt}) and the distance (D_{pt}) between the optimal allocation (w_{kpt}^{Model}) and the firm's realized holding (w_{kpt}^{actual}). b is the boundary constraint obtained from the historical distribution of the changes in allocations. Section 3.4 describes the estimation procedures. The following tables also report the pair tests between the net return model and the other models. The p_values of the pair tests have standard errors that are clustered at the firm level. All numbers are in percentages.

Panel 1: Average difference (Δ_{pt}^{Model})

				p-values of pair tests: Net return model vs	
Boundary constraint (b)	Net return	CAPM	4Factor	CAPM	4Factor
40 (75pct)	5.36	5.45	5.54	0.04	0.01
20 (50pct)	3.56	3.62	3.67	0.05	0.01
8 (25pct)	2.82	2.85	2.88	0.05	0.01

Panel 2: Average distance (D_{pt}^{Model})

				p-values of pair tests: Net return model vs	
Boundary constraint (b)	Net return	CAPM	4Factor	CAPM	4Factor
40 (75pct)	17.55	17.85	18.12	0.01	0.01
20 (50pct)	11.66	11.85	11.99	0.01	0.01
8 (25pct)	9.16	9.24	9.30	0.01	0.01

Table 6
Flow-performance relation at the aggregate level

This table reports results from regressions of aggregated flows at the fund level in year $t+1$ on fund performance and fund characteristics in year t. The aggregated flow of new money to fund f from all 401(k) plans in year t is defined as follows:

$$AGGflow_{ft} = \frac{V_{ft} - V_{f,t-1}(1 + R_{ft})}{V_{f,t-1}},$$

where $V_{ft} = \sum_{p \in \Omega_{ft}} V_{pft}$ is the total investment from all firms in year t, and R_{ft} is fund return. Fund performance and fund characteristics are described in Table 2. Standard errors are two-way clustered at the fund and year levels. The t-statistics are in brackets. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	(1) $AGGflow_{t+1}$	(2) $AGGflow_{t+1}$	(3) $AGGflow_{t+1}$	(4) $AGGflow_{t+1}$	(5) $AGGflow_{t+1}$
$Net\ return_t$	0.226**				0.052
	[2.76]				[0.47]
α_t^{CAPM}		0.438**			0.516***
		[2.78]			[2.85]
$\alpha_t^{4Factor}$			0.327**		-0.127
			[2.30]		[-0.76]
$MStar\ return_t$				0.015	-0.009
				[0.88]	[-0.68]
Controls	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Observations	21,642	21,642	21,642	21,642	21,642
Adjusted R^2	0.02	0.02	0.02	0.02	0.02

Table 7

Flow-performance relation at the firm-fund level by employee savings

This table reports results from regressions of flows at the firm-fund level in year $t+1$ on fund performance, plan and fund characteristics in year t. Fund performance, plan and fund characteristics are described in Table 2. For each year, we create quartiles of firm's average 401(k) investment per employee. *Low* and *High* contain firms with the lowest and highest investments per employee, respectively. Standard errors are clustered at the firm and year levels. The t statistics are shown in brackets. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

401(k) saving/employee	Low	2	3	High
	$Flow_{t+1}$	$Flow_{t+1}$	$Flow_{t+1}$	$Flow_{t+1}$
$Net\ return_t$	0.021**	0.024***	0.018**	0.009
	[2.41]	[3.18]	[2.76]	[1.51]
α_t^{CAPM}	0.002	0.004	0.010	0.032***
	[0.19]	[0.56]	[1.13]	[3.26]
$\alpha_t^{4Factor}$	-0.012	-0.008	-0.014*	-0.019*
	[-1.06]	[-0.89]	[-1.94]	[-1.75]
$MStar\ return_t$ ($\times 0.1$)	-0.001	0.010	0.007	-0.001
	[-0.14]	[1.36]	[1.52]	[-0.03]
Controls	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	26,481	27,253	31,079	33,650
Adjusted R^2	0.30	0.25	0.20	0.15

Table 8
Plan and employee characteristics

This table reports the average values of various characteristics of 401(k) plans in each subsample. Firms are sorted by the firm's average 401(k) investment per employee, which are described in Table 7. *Contribution cap* is the maximum amount that a firm is willing to contribute as a percentage of employee annual compensation. *Matching ratio* is the ratio of the employer's contribution to the employee's contribution. The table also reports the difference in mean tests for all variables between the highest-savings group (High) and the others (Exclude High), except for *Savings per capital*. The standard errors are clustered at the year level, and *** indicates significance at the 1% level.

401(k) saving/employee	Low	2	3	High	Exclude High
Savings per capita:					
- include company stock	4,483	14,529	27,791	121,616	15,592
- exclude company stock	3,255	10,389	19,475	51,441	11,034
Plan size (in millions):					
- include company stock	128	170	304	840	201***
- exclude company stock	91	111	189	370	130***
Number of employee	35,753	11,963	10,878	10,597	19,545***
Firm matching contribution policy:					
- contribution cap (in %)	5.22	5.41	5.41	5.48	5.35***
- matching ratio	0.59	0.62	0.66	0.73	0.62***
Firm age	24.3	22.8	25.0	33.4	24.0***

Table 9
Determinants of employee saving

This table reports results from regressions of the logarithm of the firm's average 401(k) saving per employee on plan, employee, and firm characteristics. *Firm matching ratio* and *Firm contribution cap* are the transformations of the matching ratio and the contribution cap, where the transformation is $f(x) = -e^{-x}$. The matching ratio and the contribution cap are described in Table 8. *Firm age* is the logarithm of firm age. *Employee salary* is the logarithm of an average salary of all workers who live within 25 miles of the firm's location. *Employer contribution rate* is the logarithm of the average pension contribution rate of all employers located within the same ZIP code. *Firm's debt-to-asset* is the firm's total debts to total assets. *Firm's book-to-market* is a book equity value to the market equity value. *Firm's market value* is the logarithm of market equity value. *Firm's profitability* is the gross margin to a total asset. Standard errors are clustered at the firm and year levels. *, **, and *** indicate significance at the 10%, 5%, and 1% levels, respectively.

	Employee savings	
	Coefficients	t_stat
Firm matching ratio	1.13***	[4.67]
Firm contribution cap	9.11***	[3.68]
Firm age	0.23***	[4.33]
Plan return	1.12***	[4.14]
Employee salary	0.49*	[2.06]
Employer contribution rate	0.29	[1.16]
Finance test (% of correction)	3.36**	[2.53]
Firm's debt-to-asset	0.01	[0.06]
Firm's book-to-market	−0.05**	[−2.65]
Firm's market value	−0.01	[−0.09]
Firm's profitability	−1.22***	[−7.93]
Year FE	Yes	
Observations	10,549	
Adjusted R^2	0.18	

Table A1
Employee age

Panel 1 reports average ages of employees who invested in target date funds. The table also reports the difference in mean tests between the highest-savings group (High) and the others (Exclude High), and *** indicates significance at the 1% level. *Fund weight* is the fraction of investment in fund f within 401(k) plan p in year t. β_{MKT}, β_{SMB}, and β_{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.

Panel 1: Average age.

401(k) saving/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]
β_{MKT}	−5.06***	[−14.01]
β_{SMB}	−0.94**	[−2.15]
β_{HML}	−0.04	[−0.08]
Firm FE	Yes	
Year FE	Yes	
Observations	16,179	
Pseudo R^2	0.94	