Shifting From Active to Passive: How Retirement Plans Impact Equity Prices*

March 13, 2025

ABSTRACT

Defined contribution plans have become the dominant form of retirement savings in the United States. Concurrently, there has been a notable shift toward passive investment strategies. This study investigates how the reallocation from active to passive strategies within 401(k) plans impacts equity prices. Using granular data on 401(k) fund offerings, we find that the transition to passive investing leads to aggregate price changes of around 30%, with stocks overweighted by active funds declining due to both direct outflows and reduced stock-level 401(k) ownership, a key determinant of investor demand. Our findings highlight the equity market implications of changes in pension plan investment behavior.

Keywords: Passive Investing; defined contribution plans; 401(k) plans; demand based asset pricing; price impact.

"The 401(k) has reached a tipping point in its takeover of American retirement (...) It took nearly 50 years, but half of private-sector workers are saving in 401(k)s for the first time." — February 5, 2025

The Wall Street Journal

1 Introduction

Defined contribution plans, particularly 401(k) accounts, have become the primary retirement savings vehicle for most working Americans. At the same time, a notable shift is occurring in the broader equity market, with an increasing preference for passive over actively managed investment strategies. This trend is especially relevant for retirement savings, where long-term decision-making and the challenge of beating passive indexing benchmarks play a crucial role. As more investors move toward passive strategies, it raises critical questions about whether a similar shift is occurring within defined contribution plans and what potential impact this could have on asset prices and aggregate market dynamics. This paper leverages novel data on 401(k) plan-level fund offerings, alongside 13F holdings and detailed Morningstar individual fund data, to investigate the implications of this transition within 401(k) plans.

The growing preference for passive investing among 401(k) participants is evident in the data. In 2020, index-tracking (i.e., passive) equity investments accounted for 36.55% of all equity investments in 401(k) plans, up from just 19.46% in 2007. In terms of growth, the assets held in passive funds within 401(k) plans grew by an impressive 404% from 2007 to 2020, whereas those in actively managed funds increased by only 110% over the same period.[1] While these trends are indicative of a fundamental shift in investment behavior, understanding how the reallocation from active to passive investment stratgegies within 401(k) plans affects equity prices requires further analysis of portfolio decisions—–both by mutual fund managers as well as other investors responding to this redistribution of assets.

To address this question, we draw on a recent and influential literature that employs a demand-based perspective (Koijen and Yogo, 2019) to examine how capital flows across

[1] See Figure 2 and detailed discussion in Section 2.2.

investors—or within groups of investors—impact asset prices (Haddad et al., 2025; Koijen et al., 2024). However, our approach differs in two fundamental ways.

First, our experiment contrasts with those in Koijen and Yogo (2019) and Haddad et al. (2025). The former computes counterfactual equity prices in 2016 assuming the wealth distribution among institutional investors remained at 2007 levels, thereby undoing the decline in active share of institutional investors (from 38.9% in Q4 2007 to 32.8% in Q4 2016). The latter analyzes the effects of changes in active share—using a closed-form expression—calibrated on a comprehensive set of passive investors, defined by their near-zero fixed elasticity. Importantly, both studies rely on institutional filings of a 13F form, which does not distinguish between active and passive funds within the same family. In contrast, we examine the impact of a shift from active to passive funds within 401(k) plans—originating from a subset of investors (plan sponsors) who, perhaps after consulting finance professors, realize that active strategies may be suboptimal, potentially reflecting the sizable welfare costs associated with lack of diversification (see Bhamra and Uppal, 2019, for an estimate of these costs).[2]

This detailed plan-level analysis is made possible by our new data. In particular, we are the first to leverage the novel BrightScope dataset in the context of demand-based asset pricing, estimating a demand-based system using data on individual funds (e.g., Vanguard Small-Cap Equity Fund) rather than aggregated fund family data (e.g., Vanguard). This approach is made possible through the comprehensive merging of datasets from BrightScope, Morningstar, and Thomson S34 Holdings. More broadly, our counterfactuals hold relevance both in terms of market scale (e.g., active funds managing 401(k) assets constitute approximately 90% of the total active mutual fund industry) and alignment with real-world trends (e.g., the ongoing shift of 401(k) assets from active to passive funds).

The second departure from the workhorse demand system concern its treatment of firm characteristics. While the framework in Koijen and Yogo (2019) computes counter-

[2]DC asset allocations reflect joint decisions by plan participants and sponsors. However, participants often exhibit inertia, making it more likely that the shift from active to passive is driven by plan sponsors adjusting menus to meet fiduciary duties. Plan sponsors monitor investment options and may replace underperforming funds with those showing stronger prior performance (Sialm et al., 2015).

factual prices following a given reallocation of AUM, assuming that demand functions and characteristics (other than market equity) remain unchanged, we demonstrate that the demand function of funds is affected by the 401(k) ownership of a stock, a characteristic that changes endogenously with the AUM reallocation. As capital flows from active to passive funds within a plan, the 401(k) ownership of a stock changes, creating a reinforcing effect for stocks predominantly owned by passive funds. These stocks appreciate not only because passive funds attract more capital (the classic flow-induced proportional trading effect) but also due to increased demand from other funds, as their 401(k) ownership rises. Conversely, stocks primarily owned by active funds experience a negative reinforcing effect. We show that this feedback effect is quantitatively significant, with a multiplier of 7 when the feedback is ignored, compared to 12 when the feedback is incorporated. To account for the effect of stock-level 401(k) ownership, we develop an innovative methodology for computing counterfactuals within a demand-based asset pricing framework, where *both* prices and stock characteristics change endogenously. More broadly, this methodology can be applied to scenarios where other characteristics, such as cross-ownership, play a significant role.

As discussed, our demand system framework is distinctive because it introduces a novel variable, stock-level 401(k) ownership, as a key determinant in investors' asset allocation decisions. Specifically, we show that funds consider this characteristic alongside traditional ones, such as market-to-book, when making equity allocation decisions. The influence of 401(k) ownership on investor demand is intuitive: fund managers and other investors may factor in 401(k) ownership when deciding how many shares to purchase, as they seek to invest in stocks associated with stable, long-term investors like pension funds.

Quantitately, we find that the amount of company shares owned by 401(k) plans is a key determinant – in fact, the most important one together with size – in explaining the demand of mutual funds and ETFs for a specific stock. In response to a one standard deviation increase in 401(k) stock ownership, the average active mutual fund increases its demand for the individual stock by approximately 54% (t-stat: 5.65), that translates to a change in portfolio weight from, e.g., 3.0% to 4.6%. Similarly, the average active ETF also increases its exposure to the stock by 12% (t-stat: 2.59) for the same one standard deviation change in 401(k) stock ownership.

To further validate the direct impact of 401(k) ownership on individual stocks, we apply a matching procedure that pairs stocks with similar fundamental characteristics but differing levels (e.g., high versus low) of 401(k) ownership. Our results indicate that stocks with positive 401(k) ownership achieve annual returns that are 3% to 5% higher compared to similar stocks, in terms of characteristics and investor structure, that are not held by 401(k) plans. Additionally, we examine the effect of 401(k) demand on individual stock returns through the granular instrumental variable methodology developed by Gabaix and Koijen (2023). We find that a 10% increase in the instrumented stock demand of 401(k) plans leads to an average stock price rise of 3.6%, after controlling for standard firm-specific determinants of stock returns.

Importantly, stock-level 401(k) ownership is *distinct* from other forms of institutional ownership, such as total mutual fund (Chen et al., 2000) or largest (top 10) investors' ownership of a stock (Ben-David et al., 2021). After controlling for these alternative types of ownership, the magnitude of the coefficient on our stock-level 401(k) ownership is barely affected, and so is its statistical significance. These results underscore the unique information content of stock-level 401(k) ownership for fund managers' decision.

Next we address the policy-relevant question of the price impact resulting from a shift from active to passive funds within 401(k) offerings. Given the importance of stock-level 401(k) ownership for fund managers' investment decisions, we consider two scenarios. In the benchmark scenario, where neither funds nor other investors show any preference for stock-level 401(k) ownership ($\beta_i^{IO^{401k}} = 0$ for all i), we isolate the impact of flows without introducing the feedback effect resulting from a preference for 401(k) ownership. In the second scenario, where funds' demand exhibits a positive loading on 401(k) ownership ($\beta_{funds}^{IO^{401k}} > 0$), the counterfactual prices are computed by accounting for the endogenous changes in both market equity and stock-level 401(k) ownership. This approach captures the feedback loop in which price changes arise not only from investor flow-induced trading but also from investors' positive sensitivity to stock-level 401(k) ownership.

We first observe that the value-weighted absolute percent change in equity prices, that we label equity repricing, is substantial, around 30% in both scenarios. However, this overall repricing masks important cross-sectional heterogeneity as well as differences between the two scenarios. Specifically, stocks that are overweighted by active funds within 401(k) plans experience notable price declines following a shift from active to passive funds within 401(k) offerings. These declines result from both outflows from active funds (direct effects) and reductions in stock-level 401(k) ownership (indirect effects). To better understand this result, we introduce a new metric for net flows in the counterfactual scenarios: net expected flows (NEF). This metric proxies the expected inflows or outflows from reallocating 401(k) assets from active to passive funds, capturing whether a stock was previously overweighted by active funds, suggesting that a shift toward passive funds would lead to outflows.

We find that stocks facing outflows experience an average price decline of 10% when investors do not consider stock-level 401(k) ownership (direct flow effect). However, when investors account for this characteristic, the decline intensifies to 20% reflecting the combined direct and indirect effects. In contrast, stocks with net inflows appreciate by 21% in the baseline scenario (direct flow effect), with the increase rising to 35% when preferences for stock-level 401(k) ownership are incorporated (direct and indirect effects combined). Based on this analysis, we estimate a price-flow multiplier of 12–14, exceeding existing estimates from Gabaix and Koijen (2023). This finding underscores how the growing prominence of passive investing, driven by the shift in 401(k) plan offerings, can induce substantial price effects due to the equity market's increasing inelasticity.

The reallocation of 401(k) assets from active to passive funds also has heterogeneous impacts across stocks along characteristics such as size and book-to-market ratios. Large-cap stocks generally benefit from this shift, primarily due to positive NEF values associated with inflows. Similarly, value stocks experience stronger price gains compared to growth stocks. Our feedback mechanism, driven by investors with strong preferences for stock-level 401(k) ownership, further amplifies these pricing effects. Notably, stocks with substantial inflows and increases in 401(k) ownership experience heightened price appreciation. Quantitatively, we find that a one standard deviation increase in stock-level 401(k) ownership corresponds to a 9.7% increase in a firm's market-to-book equity ratio, positively impacting counterfactual returns.

To conclude, our counterfactual analysis estimates scenarios where 401(k) plans move away from active funds and begin offering only passive funds in their investment menus, analyzing the resulting changes in asset prices. This shift not only directly affects asset prices but also alters stock-level 401(k) ownership. Since the demand function depends on stock-level 401(k) ownership, this introduces a new mechanism—one that has not been explored in the existing literature—that impacts asset prices. Our analysis underscores the transformative impact of reallocating 401(k) assets from active to passive funds on equity prices, demonstrating how these reallocations contribute to increased market inelasticity and lead to significant repricing effects across various stock categories. Furthermore, it highlights the critical role of stock-level 401(k) ownership and stresses the importance of accounting for both direct flow-driven price pressures and indirect effects from changes in stock-level ownership when assessing the broader implications of the ongoing transition from active to passive investing.

Our paper contributes to the emerging demand-based asset pricing literature. Koijen and Yogo (2019) develop a demand system approach, showing that latent demand factors explain 81% of the cross-sectional variance in stock returns. Other studies apply this framework to corporate bonds (Bretscher et al., 2021) and high-net-worth investors (Gabaix et al., 2022). Koijen et al. (2024) explore the impact of market trends, such as the shift from active to passive investing and increased demand for green firms, on price informativeness. Haddad et al. (2025) find that passive investing has led to more inelastic demand curves for individual stocks. Ben-David et al. (2023) further explore how institutional frictions in demand affect return predictability. Our contribution is to examine a specific scenario—where 401(k) plans shift from active to passive funds—and emphasize the importance of stock-level ownership in driving fund managers' investment decisions.

We also build on the literature on risk preferences and fund manager behavior. Christoffersen and Simutin (2017) show that funds managing large pension assets tend to increase exposure to high-beta stocks. Differently from their study, we estimate stock demand as a function of 401(k) plan ownership, controlling for other stock characteristics.

Lastly, our paper also contributes to the literature on pension plans. Sialm and Starks

(2012) study the investment strategies and performance of funds held primarily by retirement accounts versus those held by taxable investors, finding no significant performance differences across tax clienteles. In contrast, we show that a fund's pension asset share significantly influences its performance and asset selection. Sialm et al. (2015) analyze pension plan investment menus, finding that flows from defined contribution (DC) assets are less sticky and more sensitive to fund performance than non-DC flows due to plan sponsors adjusting investment options. However, they also document participant inertia, with limited reaction to past fund performance. Pool et al. (2016) study whether mutual fund families acting as service providers in 401(k) plans display favoritism toward their own affiliated funds, and find that fund deletions and additions are less sensitive to prior performance for affiliated than unaffiliated funds. Unlike these studies, we directly estimate individual stock demand from funds offered and not offered in 401(k) plans using a demand-based framework. Our quantification of the price and return impact of 401(k) ownership at the stock level represents a unique contribution to this literature. Moreover, whereas Sialm et al. (2015) and Christoffersen and Simutin (2017) rely on survey data about DC assets from "Pensions & Investments" (P&I) administered to domestic equity funds, we instead observe the actual 401(k) plan holdings using a novel dataset, Brightscope. Using the same data, Egan et al. (2021) address a different research question, and document heterogeneity in investment behavior of 401(k) participants, showing that higher income and more educated individuals tend to have higher equity exposure, whereas retirees and minorities tend to have lower equity exposure.

The remainder of the paper is structured as follows. Section 2 outlines the institutional framework, describes the data, and provides evidence of a shift from active to passive investing within 401(k) plans. Section 3 presents our demand-based framework and its incorporation of stock-level 401(k) ownership. Section 4 analyzes the policy-relevant question of the price impact resulting from a shift from active to passive funds within 401(k) offerings. Section 5 concludes.

2 Data

2.1 Data Sources

Our 401(k) plan holdings data comes from BrightScope Beacon, which provides comprehensive plan-level holdings data gathered from audited Form 5500 filings of private-sector defined contribution (DC) plans. This paper focuses exclusively on 401(k) plans.[3] BrightScope reports annual data on the investment options (e.g., mutual funds) available to plan participants together with the total dollar amount invested in each option. In other words, for each 401(k) plan, we observe its asset allocation on equity mutual funds (including ETFs), allocation funds (including TDFs), bond mutual funds and other types of assets (e.g., trusts and common stocks), over time. The dataset covers 708,929 different 401(k) plans over the period 2007-2020, resulting in more than 8 million fund-by-plan-by-year observations. In addition, data on fund names, fees, and tickers is also available.

Mutual fund holdings and characteristics, such as their expense ratio, category, fund domicile, investment type (e.g., ETF flag), AUM, and tickers, are obtained from Morningstar Direct.[4] We match mutual funds in 401(k) plans with Morningstar by fund tickers and names.[5]

Given our interest in the impact of 401(k) plans on US stocks, we focus on domestic equity mutual funds. Specifically, we keep mutual funds with equity ratios greater than 0.75 and remove non-US equity funds based on the Morningstar fund domicile variable. We also require funds to have at least 3 years of holdings data. While target-date funds also invest in mutual funds and ETFs, their allocation between equities and bonds follows a mechanical rebalancing process based on fund age. Therefore, we focus exclusively on equity mutual funds and ETFs directly owned by 401(k) plans. Internet Appendix A

[3]BrightScope Beacon also provides holdings for 403(b) plans, although their total market value is small relative to that of 401(k) plans.

[4]Morningstar provides exhaustive mutual fund holdings compared to other mutual fund holding databases, such as CRSP. Schwarz and Potter (2016) find that CRSP misses many SEC mandated portfolios available in SEC filings.

[5]More precisely, we map mutual fund tickers in BrightScope Beacon to Morningstar mutual fund ID (variable: *fundid*) when tickers are available in both datasets. When fund tickers are missing in either dataset, we match mutual funds by their names. We match 98.2% of mutual fund allocation in retirement plans, or a total of 3,182 mutual funds and ETFs.

describes the data cleaning procedures in detail. Our final dataset comprises a total of 2,156 funds, split between 1,763 mutual funds and 393 ETFs.

We supplement the Morningstar holdings data with stock data from CRSP and Compustat. Our empirical analysis employs the same stock characteristic as in Koijen and Yogo (2019), namely, log book equity, profitability, investment, dividends-to-book equity and market beta. Profitability is defined as operating profits scaled by book value of equity, investment as the annual growth rate of total assets, and dividends-to-book equity as the ratio of annual dividends to book equity. Market beta is estimated using a 60-month rolling regression of monthly stock excess returns (over the 1-month Treasury-bill rate) on market excess returns, requiring at least 20 months of non-missing observations.

Lastly, when analyzing the general equilibrium implications of a shift in 401(k) offerings toward passive investing in Section 4.2, we merge detailed Morningstar fund data with S34 aggregate holdings. To avoid double-counting—since Morningstar holdings are also captured in the S34 data—we adjust both the assets under management (AUM) and the holdings of S34 institutions that own mutual funds reported in Morningstar. The details of this merging procedure are provided in Internet Appendix B.

2.2 Descriptive Statistics

Figure 1 shows the distribution of 401(k) plan investments across different categories These include direct ownership of individual stocks, separate accounts, guaranteed investment contracts (GIC),[6] mutual funds (including ETFs) and collective investment trusts (CIT).

Collective investment trusts (CIT), the second largest component, account for an average of 24% of 401(k) assets under management over our sample period. These pooled investment vehicles, established by banks or trust companies, are available only to defined-contribution (DC) plan participants when the CITs are included in the DC plan menu. The Goldman Sachs Core Plus Fixed Income (bonds) and T. Rowe Price Blue Chip Growth

[6] GICs are agreements between an investor and an insurance company, typically available in retirement plans, whereby the insurance company guarantees the investor a certain rate of return in exchange for holding the deposit for a fixed period of time.

Trust (equity) are two examples of CITs offered by large financial companies to DC plan sponsors. Since CITs are not required to publicly disclose holdings, we exclude them from our analysis.[7]

The mutual fund category, which includes ETFs, is the largest component, averaging 43% of the total 401(k) assets. Figure 2 breaks this category into five groups: US equity ETFs, US equity mutual funds, US index funds, allocation funds, and others. Allocation funds consist of target-date funds and balanced funds that invest in a mix of equity and fixed income assets, while the "Others" category includes international mutual funds, bond mutual funds, money market mutual funds, and alternative investment funds. US index funds encompass both mutual funds and ETFs that are index-tracking.[8]

Our focus is on equity active funds—active mutual funds and ETFs investing in US equities—as well as passive funds. We observe a steady increase in assets held in mutual funds (orange bar) and ETFs (green bar) over time. By 2020, mutual fund assets in 401(k) plans totaled approximately $0.63 trillion, while ETF assets reached $32 billion. Although active ETF assets within 401(k) plans remain relatively small, they have growth rapidly, with ETF investments by 401(k) plans increasing at an annual rate of 16% over the past five years. More importantly, the shift toward passive investing among 401(k) participants is evident in the data. In 2020, index-tracking (i.e., passive) equity investments accounted for 36.55% of all equity investments in 401(k) plans, nearly doubling from 19.46% in 2007. Over this period, passive fund assets in 401(k) plans surged 404%, compared to a 110% increase for actively managed fund assets. These patterns underscore the growing dominance of passive strategies within 401(k) plans, reshaping the composition of retirement investments.

Panel A in Table 1 reports the cross-sectional distribution, across years, of some 401(k) plan characteristics. The first variable, fund-level IO^{401k}, measures the fraction of assets

[7] This results in a lower bound for our stock-level 401(k) ownership variable introduced in Section 3.1.

[8] We define index-tracking ETFs as large cap ETFs that track the S&P500 index (based on CRSP Objective Code: EDCL (S&P 500 Index Objective Funds). We define index mutual funds according to Morningstar (index funds and enhanced index categories). Chau et al. (2025) categorize actively managed ETFs based on fund prospectuses and find that active ETFs (mutual funds) have average expense ratios of 0.71% (0.92%). Our categorization of passive vs. active funds obtains similar results based on expense ratios, with passive funds displaying average total expense ratios of 0.10%.

under management of individual funds collectively owned by 401(k) plans. On average, 401(k) plans hold approximately 8% of fund assets, making them among the largest institutional fund investors.[9] This number can also be backed out from Figure 2, which shows that 401(k) plans' investment in US equity funds, both active and indexed, is around $1trillion in 2020, consistent with the US equity funds total assets under management of around $14 trillions.[10] Importantly, 401(k) plans invest substantially in both index and active funds, including mutual funds and ETFs. Their average ownership is 9.34% for index funds and 7.91% for active funds (second and third rows). While they are particularly large investors in active mutual funds, holding an average of 9.25% of assets (fourth row), their ownership of active ETFs remains limited at just 0.89% (fifth row). Beyond aggregate ownership, the dollar amount invested by a given 401(k) plan in a specific fund, as a fraction of the total plan assets, is quite persistent. When looking at the top 25% of the plan-fund distribution, we observe an annual autoregressive coefficient of 0.82 (sixth row). The last row reports the distribution of the stock-level $IO^{401k}(n)$ variable, which measures the fraction of a company's market capitalization collectively owned by 401(k) plans. The median ownership by 401(k) plans is 2.8%, while the 75th percentile exceeds 4%, indicating that a significant share of certain firms' equity is held within these retirement plans.

Panel B presents the dollar allocation of 401(k) plans over time to index funds, U.S. equity active mutual funds and ETFs. Over our sample period, 401(k) plans experienced a substantial increase in assets allocated to both index funds and active mutual funds. The allocation to US equity index funds grew more than tenfold, from $32.66 billion in 2008 to $380.84 billion in 2020. This reflects the increasing shift toward passive investing within retirement plans. Similarly, assets in US equity active mutual funds also saw significant growth, rising from $117.31 billion in 2008 to $628.60 billion in 2020. While passive investing has expanded rapidly, active mutual funds have remained a dominant component of 401(k) portfolios, with total allocations increasing by more than $500 billion

[9] Figure C.1 illustrates the cross-sectional distribution of fund-level 401(k) ownership, IO^{401k}, over time. The 25th and 50th percentiles exhibit minimal volatility, while the 75th percentile ranges between 6% and 12%.

[10] Specifically, in 2020, 1.04 trillion of 401(k) assets was invested in US funds: 0.63 trillion in US equity mutual funds (orange bar), 0.03 in equity ETFs (green bar), and 0.38 in US equity index funds (red bar).

over the period.

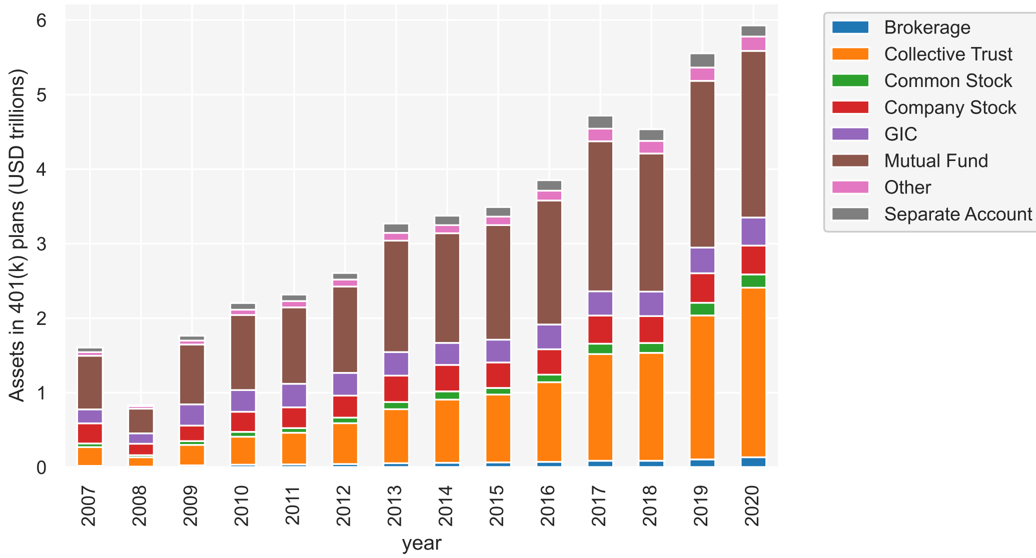


Figure 1: 401(k) plan assets. This figure shows the distribution of 401(k) plan assets into the various investment options, over time. Annual data, from 2007 to 2020.

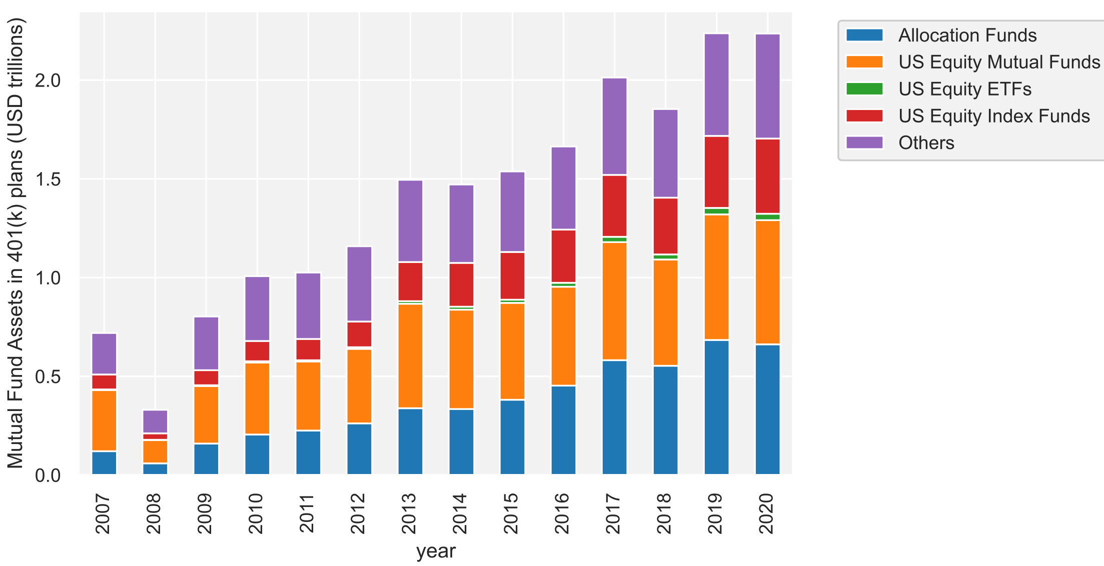


Figure 2: Distribution of assets across fund types. This figure plots the value of 401(k) investments split into various subgroups. Allocation funds are balanced funds investing in a mix of fixed income assets and equities depending on their objective, e.g., target-date funds. US equity mutual funds and ETFs include all active domestic equity funds. US equity index funds include both mutual funds and ETFs that are index-tracking. The category "Others" includes bond mutual funds, international equity mutual funds, money market funds and alternative investment funds.

3 Estimating the Impact of 401(k) Plans on Equity Demand

As discussed in Section 2, 401(k) plans allocate a substantial portion of assets to equity mutual funds and ETFs. In turn, the demand for individual stocks by mutual funds and ETFs can be influenced by 401(k) plans.

One plausible economic channel through which retirement account allocations influence the demand for individual stocks by mutual funds and ETFs is the signaling effect of significant 401(k) ownership. Fund managers may be more inclined to invest in a stock if they perceive that a large portion of its ownership is held by 401(k) plans, as this may signal a stable investor base. This is plausible because 401(k) allocations tend to be stable: 401(k) plans exhibit persistent fund allocations in terms of assets under management, as noted in Section 2, and individual fund allocations generally do not fluctuate drastically over time.[11] The fraction of a stock held by 401(k) plans can thus be treated as a stock characteristic, similar to metrics like book-to-market. Fund managers may incorporate this information when deciding how many shares of a company to purchase.[12]

Next, we formally describe our variable of interest and discuss how we adapt the asset demand framework of Koijen and Yogo (2019) to incorporate this channel.

3.1 Variable Definitions

We define *stock-level* 401(k) ownership, $IO_t^{401k}(n)$, as

$$IO_t^{401k}(n) = \frac{\sum_{j\in\{\mathcal{A},\mathcal{P}\}} \underbrace{\frac{\sum_{p=1}^{M} AUM_{p,j,t}}{AUM_{j,t}}}_{IO_{j,t}^{401k}} \times w_{j,t}(n) \times AUM_{j,t}}{ME_t(n)}$$

$$= \frac{\sum_{j\in\{\mathcal{A},\mathcal{P}\}} \left(\sum_{p=1}^{M} AUM_{p,j,t}\right) \times w_{j,t}(n)}{ME_t(n)} \tag{1}$$

[11]This is consistent with the presence of investment mandates. For example, Table 1 in Koijen and Yogo (2019) reports that 82% of stocks currently held by an institution were also held in the previous quarter.

[12]Institutional ownership of a stock, including 401(k) plan ownership, is typically accessible through public filings or third-party data providers.

where M denotes the total number of 401(k) plans investing in fund j in year t, $AUM_{p,j,t}$ the the dollar amount invested by pension plan p in fund j at the end of year t, $AUM_{j,t}$ denotes the assets under management (size) of fund j, $w_{j,t}(n)$ denotes the portfolio weight of equity fund j in stock n at the end of year t, $ME_t(n)$ is the market value of stock n, $\mathcal{A}$ denotes the set of active funds, and $\mathcal{P}$ the set of passive funds. In words, stock-level 401(k) ownership represents the fraction of stock n indirectly owned by 401(k) plans through both mutual funds and ETFs.

As illustrated in the last row of Panel A in Table 1, the interquantile ownership of individual stocks owned by 401(k) plans is substantial, ranging from 1.7% to 4.2%. Note that we focus only on indirect 401(k) ownership (i.e., through funds) since direct ownership by 401(k) plans in individual stocks is typically negligible, and we exclude this component from our analysis. Additionally, Equation (1) emphasizes that both active ($\mathcal{A}$) and passive funds ($\mathcal{P}$) influence the stock-level 401(k) variable through their respective weights.[13] In Section 4.2, we will investigate this fact in detail to examine how reallocating pension assets from active to passive funds impacts equity prices.

3.2 Model

We extend Koijen and Yogo (2019), and define the demand curve of investor i for stock n as:

$$\frac{w_{i,t}(n)}{w_{i,t}(0)} = \exp\left\{b_{0,i,t} + \beta_{0,i} mb_t(n) + \boldsymbol{\beta}'_{\mathbf{1,i}} \mathbf{X}_t(n) + \beta_i^{IO^{401k}} IO^{401k}_{-i,t}(n)\right\} \epsilon_{i,t}(n) \tag{2}$$

where $mb_t(n)$ is the log market-to-book equity of asset n at time t, $\mathbf{X}_t(n)$ is a vector of k observed characteristics of asset n at date t, and $w_{i,t}(0)$ is the portfolio weight on the outside asset.[14] Following Koijen and Yogo (2019), we include log book equity, profitability, investment, dividend-to-book equity, and market beta as characteristics. We extend

[13] Stock-level 401(k) ownership can also be expressed in terms of fund-level 401(k) ownership, $IO^{401k}_{j,t}$, i.e., the fraction of fund j's assets under management collectively owned by all 401(k) plans at the end of year t. Table 1 provides summary statistics on this variable.

[14] The outside asset includes stocks not in the investment universe. When $w_{i,t}(n) = 0$, stock n is in the investment universe of investor i, but it is not held at time t. Thus, the characteristics-based demand in (2) accommodates zero holdings.

their model by incorporating an additional component, $IO^{401k}_{-i,t}(n)$, which captures 401(k) ownership of the individual stock n.

Note that by excluding investor i from $IO^{401k}_{t}(n)$, we are studying how the portfolio choice of fund i is influenced by the stock-level 401(k) ownership of all *other* investors, thus reducing potential endogeneity concerns.[15]

Following Koijen and Yogo (2019), we assume throughout that the stock characteristics collected in $\mathbf{X}_t(n)$ are exogenous to latent demand,

$$\mathbb{E}_t\left[\epsilon_{i,t}(n) \mid \mathbf{X}_t(n)\right] = 1\,. \tag{3}$$

We now turn to our instruments for a stock's market equity and stock-level 401(k) ownership.

3.3 Instrumenting $me_t(n)$ and $IO^{401k}_t(n)$

Latent investor demand in model (2) is likely correlated with a stock's market equity, i.e., $\mathbb{E}_t\left[\epsilon_{i,t}(n) \mid me_t(n)\right] \neq 0$, because some investors are large and their individual latent demand affects stock prices.[16] To construct an instrument for the endogenous stock market equity $me_t(n)$ in equation (2), we follow Koijen et al. (2024) and use exogenous variation in investors' investment mandates to generate exogenous variation in stock demand. Let $\mathcal{S}_{i,t}$ denote the set of stocks held by investor i in period t, and assume that any stock that investor i holds during the current year, or any of the previous 11 quarters, is part of her choice set, $\mathcal{N}_{i,t} = \cup_{k=0}^{2}\mathcal{S}_{i,t-k}$ where k is expressed in years. For investor i, if $n \notin \mathcal{N}_{i,t}$, stock n is considered part of the outside asset at time t.

We compute counterfactual market equity $\widehat{me}_{-i,t}(n)$ (i.e., the instrument) as if (other) investors held an equal-weighted portfolio of all the stocks in their investment universe, excluding the investor's own holdings:

[15] Excluding investor i from the summation also addresses the concern that a stock owned only by one fund (a quite unlikely case) drives the results.

[16] Note that market equity appears in the numerator of the log market-to-book equity in the model.

$$\widehat{me}_{-i,t}(n) = \log\left(\sum_{j\neq i} AUM_{j,t}\frac{1_{n\in\mathcal{N}_{j,t}}}{1+|\mathcal{N}_{j,t}|}\right) \tag{4}$$

where $1_{n\in\mathcal{N}_{j,t}}$ is an indicator function equal to one if the stock n belongs to investor j's choice set $\mathcal{N}_{j,t}$, $AUM_{j,t}$ denotes the dollar assets under management of investor j at time t, and $|\mathcal{N}_{j,t}|$ denotes number of stocks in an investor's choice set.[17]

Note that $\widehat{me}_{-i,t}(n) = \widehat{logME}_{-i,t}(n)$ is calculated using the holdings of all investors (e.g., banks, insurances, etc.) excluding the investor i herself *and* the household sector (as in Koijen et al. (2024)). To this end, we merge 13F data from the S34 dataset, which only contains ownership data for institutions with over $100mn in assets under management, with Morningstar fund holdings (see Internet Appendix B).

The $\widehat{me}_{-i,t}(n)$ instrument can be interpreted as the counterfactual market equity of a stock, at the market clearing price, if other investors were to hold an equal-weighted portfolio within their investment universe. Similarly to Koijen and Yogo (2019), we assume that the wealth distribution AUM_j is exogenous.

Although our analysis includes fund-by-time fixed effects, which mitigate concerns that a fund characteristic jointly influences portfolio weights and 401(k) ownership, we further address potential endogeneity by introducing an instrument for stock-level 401(k) ownership. Specifically, we instrument $IO^{401k}_{-i,t}(n)$ with

$$\widehat{IO}^{401k}_{-i,t}(n) = \frac{\sum_{j=1,j\neq i}^{I} IO^{401k}_{j,t} \times \widehat{w}_j(n) \times AUM_{j,t}}{\widehat{ME}_{-i,t}(n)} \tag{5}$$

where

$$\widehat{w}_j(n) = \frac{\mathbf{1}_j(n)}{1+\sum_{m=1}^{N}\mathbf{1}_j(m)}$$

$IO^{401k}_{j,t}$ is the fraction of 401(k) assets managed by fund j, and $\widehat{ME}_{-i,t}(n)$ is the exponential of $\widehat{me}_{-i,t}(n)$ defined in (4). Note that the instrumented stock-level 401(k) ownership for a given stock (e.g., Google) will be varying across investors, since it depends, for each investor i, on the size of all other investors. For investors other than mutual funds and

[17] Although there are $|\mathcal{N}_{j,t}|+1$ assets including the outside asset, there are only $|\mathcal{N}_{j,t}|$ degrees of freedom implied by the budget constraint, since asset weights must sum to unity.

ETFs managing 401(k) assets ($IO^{401k}_{j,t} > 0$), such as insurances, the numerator in (5) will not change, and variation across investors only occurs in the denominator $\widehat{ME}_{-i,t}(n)$. Appendix C.4 discusses the relevance of our instruments.

Koijen and Yogo (2019) highlight the importance of latent asset demand, defined as the component of the demand function unexplained by the model covariates. We conjecture that an important component of the variation in latent asset demand is attributable to the fraction of stock n owned in aggregate by 401(k) plans. Next, we estimate the magnitude of these demand effects.

4 Demand for Individual Stocks by Mutual Funds and ETFs

We first assess the significance of stock-level 401(k) ownership in driving stock demand. Like market beta, 401(k) ownership is a firm-specific characteristic that can influence the stock's demand by funds. For example, fund managers may favor stocks with high 401(k) ownership, viewing them as having a more stable investor base.

We then estimate the following panel regression:

$$\log\left(\frac{w_{i,t}(n)}{w_{i,t}(0)}\right) = b_0 + \beta_0 \widehat{mb}_{-i,t}(n) + \boldsymbol{\beta_1'}\mathbf{X}_t(n) + \beta^{IO^{401k}} \widehat{IO}^{401k}_{-i,t}(n) + \alpha_{i,t} + \tilde{u}_{i,t}(n) \quad (6)$$

where the dependent variable represents the demand of stock n by fund i at time t with respect to the outside asset, $\widehat{mb}_{-i,t}(n)$ is the instrumented log market-to-book equity of firm n at time t, $\mathbf{X}_t(n)$ is a vector of controls that includes the firm-specific characteristics specified in Koijen and Yogo (2019), $\widehat{IO}^{401k}_{-i,t}(n)$ is the instrumented fraction of stock n cumulatively owned by 401(k) plans through funds, and $\alpha_{i,t}$ are fund-by-time fixed effects. In addition to the variables used in Koijen and Yogo (2019), we also present results controlling for three alternative ownership variables that may influence fund demand for individual stocks: the fraction of a stock owned by the top ten investors (Ben-David et al., 2021), a stock's total mutual fund ownership, and the stock ownership by institutional investors categorized by different levels of portfolio turnover and diversification (Bushee, 1998).

Panel A in Table 2 shows the results from the panel regression (6) for the entire universe of funds (columns (1)-(3)), mutual funds (columns (4)-(6)), and ETFs (columns (7)-(9)). We report three-way (funds, time and stock) clustered standard errors.[18] Fund-stock observations are AUM-weighted. Furthermore, to compare regression coefficients, we standardize all variables. Across specifications, the coefficient on stock-level 401(k) ownership, $\widehat{IO}^{401k}_{-i,t}(n)$, is positive, it ranks second in terms of magnitude after size (among the characteristics included in $\mathbf{X}_t(n)$), and it is statistically significant even after controlling for well known drivers of expected returns such as market beta, book-to-market, and profitability. This result highlights the relevance of stock-level 401(k) ownership as an important characteristic for fund allocation decisions. The coefficient for the universe of funds (0.32, t-stat=4.71) is mostly determined by mutual funds. Specifically, mutual funds display a loading on stock-level 401(k) ownership of 0.54 (t-stat=5.65), which is more than four times higher than that of ETFs, at 0.12 (t-stat=2.59).[19] Controlling for stock ownership by the top ten investors (columns 2, 6, and 11), or total mutual fund ownership (columns 3, 7, and 11) of a stock, does not affect our results. In columns (4), (8), and (12) we control for the three groups of institutional investors delineated in Bushee (1998): "quasi indexed (QIX)" (institutions that are widely diversified and do not trade much); "dedicated (DED)" (institutions whose holdings are more concentrated, but do not trade much); and "transient (TRA)" (institutions whose holdings are diversified but trade often in and out from individual stocks). Also in this case, the association between stock-level 401(k) ownership and mutual fund demand for individual stocks remains positive and statistically significant. These robustness checks highlight the uniqueness and relevance of 401(k) stock-ownership with respect to other types of institutional ownership.[20]

Although, specification (6) includes fund-by-time fixed effects, mitigating concerns of

[18] Using two-way - funds and time - clustered standard errors obtains similar results.

[19] Appendix Table C.1 reports the same results without weighting the observations by the fund AUM. The coefficients for mutual funds and ETFs are 0.41 (t-stat=7.81) and 0.15 (t-stat=3.49), respectively, thus confirming a stronger effect for the former. Appendix Table C.2 reports the results without winsorizing fund size at the top 2.5% level. Our conclusions continue to hold. In fact, the estimates are even larger, with the coefficients for mutual funds and ETFs being 0.179 (t-stat=10.00) and 0.116 (t-stat=9.63), respectively. Finally, Appendix C.5 highlights the importance of our variable in explaining latent demand, defined as in Koijen and Yogo (2019). Specifically, we show that incorporating $IO^{401k}_{-i,t}(n)$ alongside Fama-French-type characteristics typically used in demand-based asset pricing reduces the standard deviation of latent demand, making portfolio weights less extreme.

[20] In Internet Appendix F.1 we verify that our stock-level 401(k) results are robust to using s34 data instead of Morningstar.

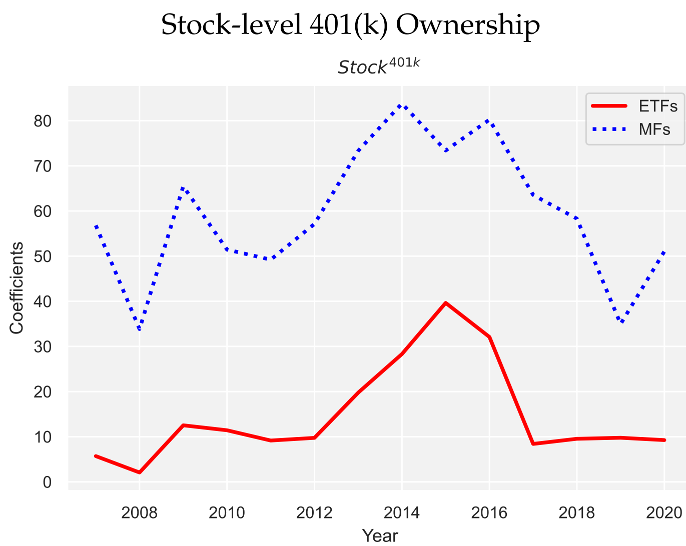


Figure 3: Coefficients on 401(k) ownership. This figure shows the annual coefficient in equations (6) on stock level 401(k) ownership, separately for mutual funds and ETFs, estimated by pooled OLS using assets under management as weights. The regression is estimated annually, and it includes fund-level fixed effect. Variables are standardized (within each year) to make coefficients comparable. We multiply the coefficients on 401(k) ownership by 100, so that they can be interpreted as the percentage change in demand per one standard deviation change in the characteristic. The sample period is from 2007 through 2020.

a fund characteristic jointly influences portfolio weights and 401(k) ownership, in Panel B of Table 2 we repeat our analysis and assess the imapct of 401(k) stock ownership on fund allocations using a sample of funds that do *not* appear on the 401(k) menus, i.e., funds that do not manage 401(k) assets. By doing so, we remove any potential selection effect arising from 401(k) plans choosing funds with specific characteristics (e.g., funds from large families) or similar investment strategies (e.g., investing in growth stocks). Note that for this particular sample of funds that do not manage any 401(k) assets, $IO_{-i,t}^{401k}(n) \equiv IO_t^{401k}(n)$. Importantly, we continue to find a large and statistically significant coefficient on $IO_{-i,t}^{401k}(n)$.[21] Overall, the evidence in Table 2 underscores the importance of 401(k) ownership in driving the funds' demand for stocks.

[21]For the sample of mutual funds that do not control pension assets, the coefficient on $IO_t^{401k}(n)$ is 0.46 (t-stat=4.47) (Table 2, Panel B). Not only this coefficient is similar in magnitude to the one obtained in our benchmark sample of fund managers controlling pension assets, but also it is robust to alternative specifications. In particular, we estimate a loading of 0.50 (t-stat=5.44) if we do not weight observations by the fund AUM (Appendix Table C.1, Panel B) and 0.099 (t-stat=3.65) when we do not winsorize fund TNA (Table C.2, Panel B). Lastly, in Table C.3, the coefficient remains sizable and statistically significant even when the stock-level 401(k) ownership enters the specification with a lag.

Figure 3 shows the evolution of the coefficient on stock-level 401(k) ownership, $IO^{401k}_{-i,t}(n)$, over time.[22] The coefficient is always larger and more volatile for mutual funds than for ETFs; in general, the magnitude of the coefficients are in line with the values reported in Table 2. Panel A of Table 3 reports GMM estimates of the main specification of the non-linear version of equation (6). The result shows a positive on stock-level 401(k) ownership of 0.25 – the second largest within the set of characteristics $\mathbf{X}_t(n)$ – statistically significant at the 1% level (t-stat: 25.49).

Next, we examine the price impact of stock-level 401(k) ownership in a reduced-form setting, before addressing our policy-relevant question of the price impact resulting from a shift from active to passive funds within 401(k) offerings using a structural equilibrium framework.

4.1 Stock Demand and Price Impact of 401(k) Plans: Reduced Form Estimation

4.1.1 Matched Sample of Low and High Stock-level 401(k) Ownership

In Section 4, we estimated the impact of 401(k) plans for individual stock demand. In this section, instead, we quantify the direct impact 401(k) ownership has on individual stock *returns* by employing a matching analysis: we compare otherwise identical stocks that only differ by 401(k) ownership, and analyze their return dynamics. In other words, we match pairs of similar stocks together, one displaying positive 401(k) ownership (the treated stock), while the other not owned by 401(k) plans (the control stock). This matching exercise allow us to evaluate whether stocks belonging to the treatment and control groups, which are otherwise identical, perform differently.

We start our analysis identifying, every year, the largest institutional investor for each stock (e.g., Blackrock, Fidelity, etc.).[23] This process yields a time series of each stock's largest institutional investor. We then count how often each investor holds the top position across all stocks and years, extracting the ten most frequent investors' names. This

[22] Figure C.2 shows the coefficients on the other covariates.

[23] Since 13F holdings are quarterly, we select the investor ranked at the top most of the quarters within a year. If there is a draw, we select the largest investor in terms of AUM.

list includes Blackrock, Vanguard, Fidelity, Dimensional Fund Advisors, among others.

Next, for each of these ten investors, we select the subset of stocks for which this investor (e.g., Vanguard) is the largest. Within this subset, we match stocks with positive 401(k) ownership (*treated* group) to comparable stocks without 401(k) ownership (*control* group). Matched stocks share the same largest investor (e.g., Vanguard), and exhibit similar (i) portfolio weights in the largest investor's portfolio; (ii) size; and (iii) book-to-market. More precisely, we sort the candidate "matching" stocks on the difference between their market capitalization and the treated stock's market capitalization. This generates a "market cap rank," where the candidate stock with rank = 1 has a market cap closest to the one of the treated stock. We apply the same ranking methodology to the book-to-market. The stock with the *smallest* sum of market cap and book-to-market ranks for each treated stock, every year, is selected as the "matched" stock in our control group.

We repeat the above matching procedure for each stock owned by all of the ten largest investors. Lastly, we estimate a panel regression of annual returns of the matched pair of stocks on a treated dummy variable and various stock-level controls. Controls include lagged size, book-to-market, beta, and momentum. Standard errors are double clustered by stock and time.

Panel A of Table 4 reports the average characteristics of the matched sample, while Panel B presents the regression results. The coefficient on the "treated dummy" is 5% in a specification without controls and around 3.2% after accounting for main determinants of cross-sectional return variation, such as beta, book-to-market, log market equity and momentum. In other words, stocks with positive 401(k) ownership tend to earn 3%-5% higher return than comparable stocks – matched on characteristics and investor structure – that are not held by 401(k) plans.

4.1.2 Granular Instrumental Variable (GIV) Approach

The previous section established a relationship between 401(k) ownership and stock returns. In this section, we employ the granular instrumental variable (GIV) approach of Gabaix and Koijen (2024) to provide more causal evidence on the role of 401(k) demand in driving individual stock returns.

Specifically, similar to Fan et al. (2022), we define the value-weighted 401(k)'s demand for individual stocks as

$$\text{Demand}_t^{401(k),VW}(n) = \sum_{i=1}^{N_t(n)} w_{i,t-1}(n) \times \frac{\text{Shares}_{i,t}(n) - \text{Shares}_{i,t-1}(n)}{\text{Shares}_{i,t-1}(n)} \tag{7}$$

where $N_t(n)$ is the total number of 401(k) plans that own stock n at time t, and the weight $w_{i,t-1}(n)$ represents the proportion of stock n owned by 401(k) plan i at the end of the preceding year $t-1$, which is calculated as the ratio of the shares of stock n held by 401(k) plan i to the total shares of stock n collectively held by all 401(k) plans:

$$w_{i,t}(n) = \frac{\text{Shares}_{i,t}(n)}{\sum_{j=1}^{N_t(n)} \text{Shares}_{j,t}(n)}$$

We also compute the corresponding equally-weighted demand as:

$$\text{Demand}_t^{401(k),EW}(n) = \frac{1}{N_t(n)} \sum_{i=1}^{N_t(n)} \frac{\text{Shares}_{i,t}(n) - \text{Shares}_{i,t-1}(n)}{\text{Shares}_{i,t-1}(n)} \tag{8}$$

To estimate the relationship between (value-weighted) demand originated by 401(k) plans, $\text{Demand}_t^{401(k),VW}(n)$, and individual stock returns, we run the following stock-level panel regression:

$$r_t(n) = \beta_0 + \beta_1(n) \times \left(\widehat{\text{Demand}_t^{401(k),VW}}(n)\right) + \varepsilon_t(n)$$

by instrumenting $\text{Demand}_t^{401(k),VW}(n)$ with the demand "shock" ($\text{Demand}_t^{401(k),VW}(n) - \text{Demand}_t^{401(k),EW}(n)$), i.e., the difference between the value-weighted and equally-weighted flows.[24]

Table 6 reports the estimation results, controlling for size, beta, and book-to-market, stock- and time (year) fixed effects. We observe that, in the most stringent specification, the coefficient on the instrumented demand is about 0.37, suggesting that for a ten percent

[24] As in standard IV setups, we first regress the endogeneous $\text{Demand}_t^{401(k),VW}(n)$ on the difference between the value- and equally-weighted demand (first stage), and use the (exogeneous) fitted value as regressor in the second IV stage.

increase in 401(k) demand, stock prices increase by 3.7%.

4.2 Price Impact: Shift from Active to Passive Funds within 401(k) plans

Over the past decade, a notable shift has occurred from active to passive funds (Koijen et al., 2024; Haddad et al., 2025), largely driven by efforts to minimize fees.[25] While existing studies examine this shift across a broad range of investors, we focus specifically on the transition occurring within defined contribution plan fund offerings. In this section, we assess how this reallocation within 401(k) plans affects equity prices.

Figure 2 illustrates the evolution of 401(k) asset allocation between active mutual funds (MF) and ETFs, and passive index funds. While assets have grown in both categories, the increase has been substantially larger for passive funds (red bar). By 2020, passive, index-tracking equity investments accounted for 36.55% of all equity investments in 401(k) plans, up from 19.46% in 2007. Over this period, the assets held in passive funds within 401(k) plans grew by an impressive 404%, whereas those in actively managed funds increased by only 110%.

This shift in asset allocation has also influenced fund-level 401(k) ownership, with a pronounced differential between passive and active funds. Figure 4 reveals that the median fund-level 401(k) ownership (solid black line) and the 75th percentile (dashed blue line) are markedly higher for passive funds, averaging 2% vs. 5% and 10% vs. 15%, respectively. This increasing preference for passive funds at the fund level has significant implications for stock-level ownership. As depicted in Figure 5, the median stock-level 401(k) ownership via passive funds has increased steadily from 0.3% to 1.1%, while stock-level ownership through active funds has remained relatively stable at around 1.5%.

Building on this evidence, we now investigate how increased 401(k) investments in passive funds influence individual stock prices. Using a novel counterfactual analysis, we assess the effects of reallocating 401(k) assets entirely from active to passive funds. Specifically, we explore the price impact on individual stocks if 401(k) plans offered only

[25] As an example, investors across Europe and the UK who bought index trackers instead of active funds have saved nearly £80bn in fees over the past 12 years, putting substantial pressure on the actively managed fund industry. See here.

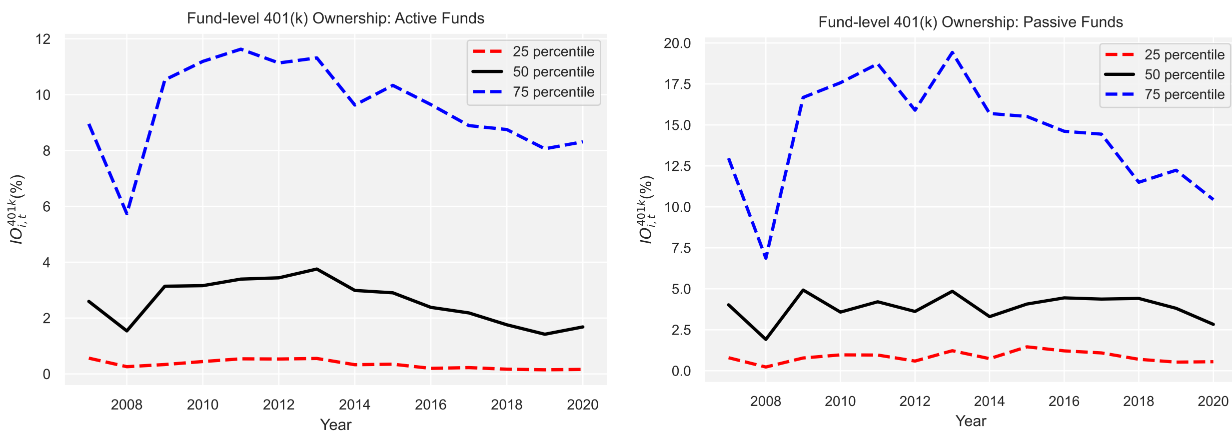


Figure 4: Active vs. passive fund-level 401(k) ownership. We extract the fund-level 401(k) ownership, $IO^{401k}_{j,t}$, from the middle equation of the expression in (1) and plot it separately for $j \in \mathcal{A}$ (active funds) and $j \in \mathcal{P}$ (passive funds). The figure displays the distribution of the resulting $IO^{401k}_{j,t}$, over time, with active funds in the left panel and passive funds in the right panel. The sample period spans from 2007 to 2020.

passive funds in their investment menus. How would such a shift reshape equity markets and affect the pricing of individual stocks?

We base our general equilibrium counterfactual analysis on the work of Koijen et al. (2024), extending it in several important ways. First, we consider a more realistic scenario where reallocating pension assets from active to passive funds reduces the size of the active management industry without entirely eliminating it. In contrast, Koijen et al. (2024, Section 8) focus on reallocating the entire assets under management (AUM) of an investor group to assess its relative importance. Second, we analyze an equilibrium where these flows affect prices both *directly*, via standard demand pressure mechanisms, and *indirectly*, by altering 401(k) stock ownership—a critical determinant of investor demand as documented in Section 4. We hypothesize that stocks heavily owned by active funds managing substantial pension assets will experience price declines following outflows (direct effect). This decline is likely magnified for stocks overweighted by active funds compared to passive funds, as reduced 401(k) stock-level ownership—now coming solely from passive funds with lower weights in those stocks in the counterfactual scenario—would further suppress demand (indirect effect).

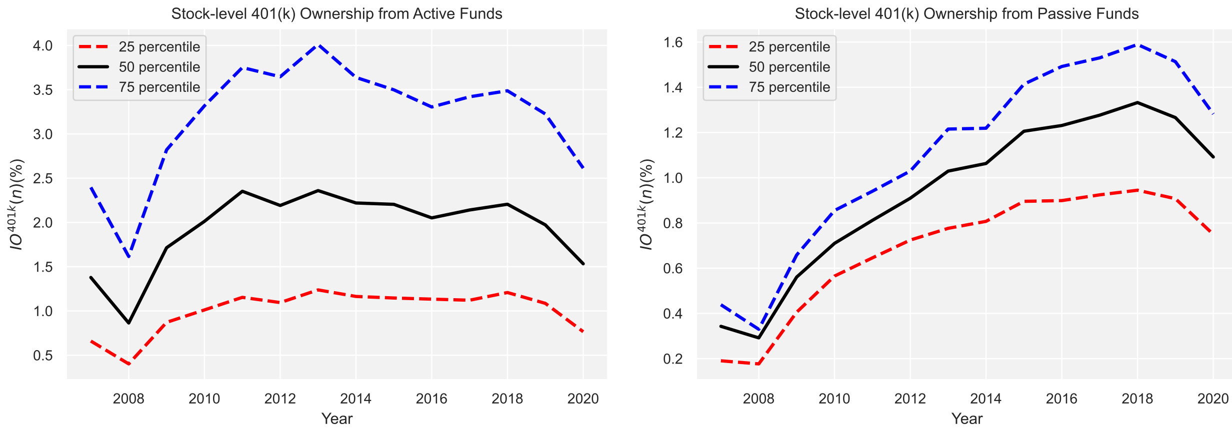


Figure 5: Active vs. passive stock-level 401(k) ownership. This figure plots the cross-sectional distribution of 401(k) stock-level ownership for active funds (left panel) and passive funds (right panel) over time. For each stock in a given year, we calculate the ratio of 401(k) investments in the stock by all mutual funds/ETFs to its market capitalization. Then, we plot the 25th, 50th, and 75th percentile of the stocks ranked by this fraction for active funds (left panel) and passive funds (right panel).

4.2.1 Counterfactual Framework

Our general equilibrium analysis relies on a comprehensive dataset of investor holdings, which combines 13F filings from Thomson-Reuters with mutual fund holdings data from Morningstar (see Internet Appendix B). This approach differs from that of Koijen and Yogo (2019) and Koijen et al. (2024), who only consider funds at the "family" (e.g., Vanguard) level. For our purposes, it is essential to distinguish between active and passive funds *within* the same fund family, as we focus on reallocations within a single group of investors (e.g., across mutual funds) rather than across different investor groups (e.g., from hedge funds to mutual funds).

We make this distinction explicit through the following notation. Let $\mathcal{A}(p)$ represent the set of active mutual funds and ETFs offered by pension plan p, and $\mathcal{P}(p)$ denote the set of passive funds offered by p.[26] We define the universes of all active and passive funds that manage pension assets as $\mathcal{A} = \bigcup_p \mathcal{A}(p)$ and $\mathcal{P} = \bigcup_p \mathcal{P}(p)$, respectively.[27] The remaining sets of investors consists of I entities indexed by $i = 1, \ldots, I$, where investor $i = 1$ represents households. The household sector, defined in this way, acts as a residual

[26] All funds offered by p are classified as either active or passive. See footnote 8 for the definition of active and passive funds.

[27] The assets under management by active funds that handle 401(k) pension money represent a substantial portion – over 90% throughout our sample period – of the total dollar assets managed by the universe of active funds in our sample. Our counterfactuals therefore accurately reflect potential real economic effects.

category that holds all remaining shares not owned by institutional investors, such as insurance companies, banks, investment advisors, and mutual funds.

In the counterfactual scenario, we reallocate the assets held by pension plan p in active funds at time t, denoted $AUM_{p,k,t}$ where $k \in \mathcal{A}(p)$, into passive funds offered by the same pension plan. For each plan p, we compute the *flows* to the passive fund j at time t generated from this reallocation as:[28]

$$F_{p,j,t} = \frac{AUM_{j,t}}{\sum_{i \in \mathcal{P}(p)} AUM_{i,t}} \times \sum_{k \in \mathcal{A}(p)} AUM_{p,k,t} \qquad \forall j \in \mathcal{P}(p) \tag{9}$$

This reallocation induces an outflow from the active fund management industry:[29]

$$F_{k,t} = -\sum_{p \in 401(\mathrm{k})} AUM_{p,k,t} \qquad \forall k \in \mathcal{A} \tag{10}$$

and a corresponding inflow into the passive management industry:

$$F_{j,t} = \sum_{p \in 401(\mathrm{k})} F_{p,j,t} \qquad \forall j \in \mathcal{P} \tag{11}$$

These flows define a new wealth distribution across institutional investors:

$$AUM_{i,t}^{CF}\left(\mathbf{P}_t^{CF}\right) = AUM_{i,t}\left(w_{i,t}(0) + \sum_{n \in \mathcal{N}_{i,t}} \frac{P_t^{CF}(n)}{P_t(n)} w_{i,t}(n)\right) + F_{i,t} \tag{12}$$

where superscript CF denotes counterfactual values of the corresponding variables, $w_{it}(n)$ and $w_{it}(0)$ denote investor i's portfolio weight on stock n and the outside asset, respectively, in period t. As in Koijen et al. (2024) we do not change the asset demand functions (i.e., the demand coefficients and latent demand) in the counterfactual. Once capital flows are specified, the counterfactual $AUM_{i,t}^{CF}$ distribution can be computed at the converged vector of equity prices using the above equation.

[28] If a 401(k) plan does not offer passive funds, we reallocate the active funds' assets to the SPDR S&P 500 ETF (SPY). In our sample, about 25.2% of plans do not offer passive funds on their menu. In terms of assets, these plans account for 13.7% of total 401(k) dollar investments. Over the last three years of the sample, this figure dropped to 16.1%, and it is expected to decline further in the future.

[29] If plan p does not own the active fund k, then $AUM_{p,k,t} = 0$.

The counterfactual equity price of firm n is determined by the market-clearing condition for each stock n:

$$P_t^{CF}(n) = \sum_{i=\mathcal{A},\mathcal{P},1...,I} AUM_{i,t}^{CF}\left(\mathbf{P}_t^{CF}\right) w_{i,t}^{CF}(n; \mathbf{P}_t^{CF}; IO_t^{401(k),CF}(n)) \tag{13}$$

Counterfactual market equity is defined as the sum of asset demand, equal to AUM multiplied by the portfolio weight, across all investors. There are, however, several unique aspects of our setting to consider. First, the summation includes all investors, including active mutual funds, which maintain a reduced but still positive AUM despite outflows. Second, equilibrium prices, along with the AUM revalued at these counterfactual prices, affect our stock-level 401(k) characteristic, which in turn affect demand functions. This relationship can be seen in the following expression:[30]

$$IO_t^{401k,CF}(n) = \frac{\sum_{j\in\mathcal{P}}\left(\sum_{p=1}^{M} AUM_{p,j,t}^{CF}\right) \times w_{j,t}^{CF}(n)}{P_t^{CF}(n)} \tag{14}$$

Crucially, the weights determining stock-level 401(k) ownership are exclusively those of passive funds. Substituting equation (14) into (13), we observe that the weight of investor i on stock n depend on the weight of passive fund j:

$$P_t^{CF}(n) = \sum_{i=\mathcal{A},\mathcal{P},1...,I} AUM_{i,t}^{CF}\left(\mathbf{P}_t^{CF}\right) w_{i,t}^{CF}(n; \mathbf{P}_t^{CF}; w_{j,t}^{CF}(n)) .$$

However, a key simplification arises from the independence of passive fund weights from stock-level 401(k) ownership. As shown by Koijen and Yogo (2019, Section C. Estimation on a Hypothetical Index Fund), an index (i.e., passive) fund's portfolio weights are given by:

$$\frac{w_j(n)}{w_j(0)} = \exp\left\{me(n) + \text{const}_j\right\} \qquad j \in \mathcal{P}$$

This ensures that $IO_t^{401k,CF}(n)$ is a function of counterfactual prices only. As a result, the weights in (14) depend solely on $P_t^{CF}(n)$, avoiding a direct circular relationship between

[30]The AUM of a passive fund $j \in \mathcal{P}$ attributed to pension assets of 401(k) plan p, transitions from $AUM_{p,j,t}$ to $AUM_{p,j,t}^{NEW} = F_{p,j,t} + AUM_{p,j,t}$ where $F_{p,j,t}$ is defined in (9). Equation (12) then gives the counterfactual AUM at the converged vector of equity prices.

weights and counterfactual 401(k) ownership. Consequently, the market-clearing equation (13) can be rewritten as:

$$P_t^{CF}(n) = \sum_{i=\mathcal{A},\mathcal{P},1\ldots,I} AUM_{i,t}^{CF}\left(\mathbf{P}_t^{CF}\right) w_{i,t}^{CF}(n;\mathbf{P}_t^{CF}) \tag{15}$$

The counterfactual equity prices can then be solved by iterating on this equation until convergence, using the algorithm provided in Koijen and Yogo (2019, Appendix C). Overall, the independence of passive fund weights from stock-level 401(k) ownership enables the transition from (13) to (15).[31] While equation (15) facilitates the algorithm's implementation, equation (13) better captures the feedback loop intrinsic to our framework, which features a "two layer" structure: 401(k) pension plans invest in mutual funds and ETFs which, in turn, invest in individual stocks. This structure demonstrates how reallocating assets from active to passive funds within a plan directly impacts equity prices, which then influences the counterfactual AUM distribution (via Eq. (12)). Both elements subsequently affect the stock-level 401(k) ownership (see Eq. (14)), which then feeds back into equilibrium prices, creating a recursive cycle, a *unique* feature of our setup.

Finally, we analyze two scenarios where investors exhibit a preference for stock-level 401(k) ownership, i.e., $\beta_i^{IO^{401k}} > 0$. In the first scenario, only funds[32] have a non-zero coefficient on 401(k) ownership, $\beta_{funds}^{IO^{401k}} > 0$, while we restrict the demand functions of all other investors to be independent of stock-level 401(k) ownership. In the second scenario, *all* investors, except households, show a preference for stock-level 401(k) ownership, $\beta_{ALL}^{IO^{401k}} > 0$.[33] We compare these two scenarios to a benchmark case in which neither funds nor other investors display any preference for stock-level 401(k) ownership, $\beta_i^{IO^{401k}} = 0$ for all i. This benchmark isolates the impact of flows without introducing the feedback effect generated by a preference for stock-level 401(k) ownership.

All the counterfactual results in the next section are estimated every year and aggregated over the sample 2010-2019. Internet Appendix D describes the pseudo-algorithm

[31] Note that the weights of investors other than passive funds still depend on counterfactual 401(k) ownership, but they do so indirectly.

[32] In the counterfactuals, we refer to active mutual funds and active ETFs collectively as funds.

[33] Figure E.4 in Internet Appendix E reports the estimated demand coefficient on 401(k) stock-level ownership for all investors other than mutual funds and ETFs.

used in estimating the counterfactual, while Internet Appendix E presents an alternative analytical approach to gauge the price impact triggered by an exogenous change in 401(k) ownership in our demand system.

4.3 Results

We first compute the value-weighted absolute percent change in equity prices, i.e., equity repricing (Koijen et al., 2024), and find it to be substantial at approximately 30% across all scenarios ($\beta_{funds}^{IO^{401k}} > 0$, $\beta_{ALL}^{IO^{401k}} > 0$ and the benchmark where no investors exhibit a preference for stock-level 401(k) ownership, $\beta_i^{IO^{401k}} = 0$). However, this overall repricing masks important cross-sectional heterogeneity and differences between the scenarios.[34]

Intuitively, if a stock is overweighted by active funds within 401(k) plans, its price is likely to depreciate in the counterfactual scenario not only due to dollar outflows from these active funds but also because of a reduction in its stock-level 401(k) ownership, which serves as an amplification mechanism. To test this conjecture, we define the following net expected flow measure (*NEF*):

$$NEF^{401k}(n) = -\left[\sum_{j\in\{\mathcal{A},\mathcal{P}\}}\left(\sum_{p=1}^{M} AUM_{p,j,t}\right)\times w_{j,t}(n) - \sum_{j\in\mathcal{P}}\left(\sum_{p=1}^{M} AUM_{p,j,t}^{NEW}\right)\times w_{j,t}(n)\right] \tag{16}$$

where the post-redistribution total assets that 401(k) plan p invests in passive fund j is given by:

$$AUM_{p,j,t}^{NEW} = F_{p,j,t} + AUM_{p,j,t} \qquad j\in\mathcal{P}$$

Recall from equation (9) that $F_{p,j,t}$ represents the flow resulting from the redistribution of 401(k) assets within plan p, reallocating assets from active funds $k\in\mathcal{A}(p)$ ($AUM_{p,k\in\mathcal{A}(p),t}$) to passive funds $j\in\mathcal{P}(p)$ available within the same plan p. Equation (16) thus measures the extent to which stock n was over- or underweighted prior to the transition, relative to the new scenario where only passive funds are offered by the pension plan. We empha-

[34]Specifically, we compute $\frac{\sum_{n=1}^{N}\left|P_t^{CF}(n)-P_t(n)\right|}{\sum_{n=1}^{N}P_t(n)}$ and find repricing effects of 34% in the benchmark case ($\beta_i^{IO^{401k}} = 0$), 35% when only funds exhibit a preference for 401(k) ownership ($\beta_{funds}^{IO^{401k}} > 0$), and 36% when all investors except households display such a preference ($\beta_{ALL}^{IO^{401k}} > 0$).

size that the NEF measure defined in Equation (16) relies on observed quantities, such as $AUM^{NEW}_{p,j,t}$, rather than counterfactual values like $AUM^{CF}_{p,j,t}$ derived from the market-clearing condition (13). As such, it serves as a gauge of net expected flows, which may differ from those obtained in the counterfactual equilibrium. Importantly, Figure D.1 demonstrates that our NEF provides a reliable proxy for actual realized flows.

To aid interpretation, we scale the $NEF^{401k}(n)$ measure for each stock n by its market capitalization, $ME(n)$. The resulting $NEF^{401k}(n)$ (%) measure can thus be interpreted as a percentage demand shock, i.e., dollar inflows or outflows relative to the stock's size. $NEF^{401k}(n)$ (%) can be negative, indicating outflows from stock n (typically when the stock was overweighted by active funds), or positive, indicating inflows. Finally, we sort stocks into quintiles based on their $NEF^{401k}(n)$ (%).

Figure 6 presents a scatter plot of counterfactual stock returns, with the y-axis showing returns under a scenario where funds exhibit a preference for 401(k) ownership, i.e., $\beta^{IO^{401k}}_{funds} > 0$, and the x-axis showing returns under a scenario where investors ignore this stock-level characteristic, i.e., $\beta^{IO^{401k}}_{i} = 0$. Stocks are color-coded based on their NEF values: gray dots represent the top quintile (inflows), green, violet, and light blue indicate quintiles 2 through 4, respectively, and black denotes quintile 1 (outflows).[35]

If all stock returns fell along the 45-degree line, it would suggest that the feedback effect from funds preferring stocks with higher 401(k) ownership has no impact. However, the observed scatter shows stocks above and below the 45-degree line. This pattern reveals that the reallocation of 401(k) assets to passive investments can create both overpricing and underpricing of stocks when fund demand functions reflect a preference for stock-level 401(k) ownership ($\beta^{IO^{401k}}_{funds} > 0$), compared to a scenario where such preferences are absent.

The pricing behavior due to the direct effect of flows is intuitive: stock appreciation becomes more pronounced as the flows a stock receives ($NEF^{401k}(n)$ (%)) increase. More importantly, we observe an indirect effect driven by changes in stock-level 401(k) owner-

[35]Under a scenario where funds exhibit a preference for 401(k) ownership, the counterfactual return is defined as the difference between the log of the counterfactual price, when mutual funds care about stock-level 401(k) ownership ($\beta^{IO^{401k}}_{funds} > 0$), and the log of the initial price, i.e., before the reallocation of AUM from active to passive funds occurs.

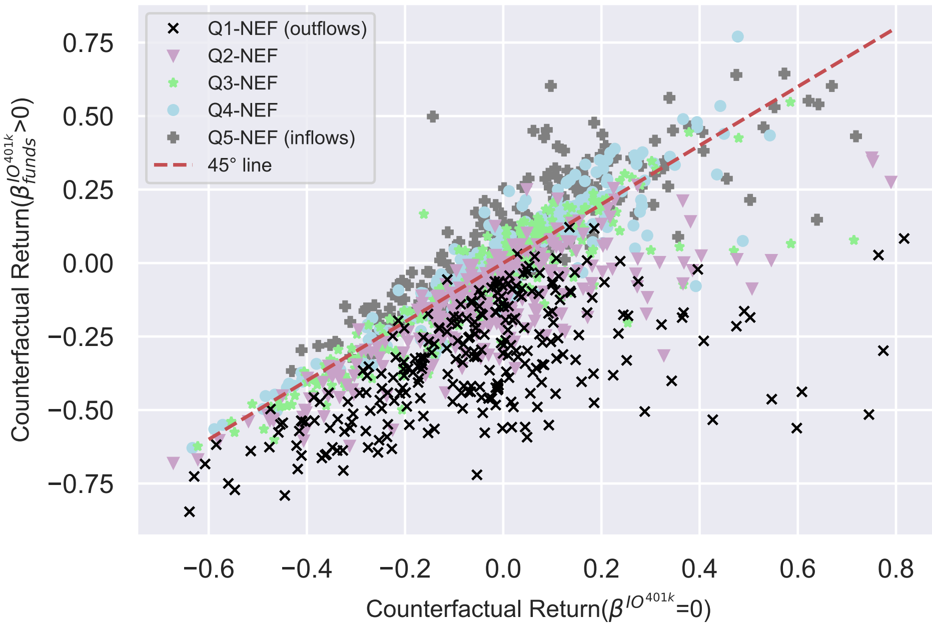


Figure 6: Counterfactual returns: impact of active funds' 401(k) ownership preferences. This figure presents a scatter plot of the counterfactual stock returns, comparing the scenario where active funds display a preference for 401(k) stock-level ownership ($\beta^{IO^{401k}} > 0$, y-axis) to the scenario where they are indifferent to this characteristic ($\beta_i^{IO^{401k}} = 0$ for all investors i, x-axis). Stocks are color-coded based on their net expected flow measure, $NEF^{401k}(n)$ (%).

ship. Stocks experiencing inflows in the counterfactual scenario see an increase in their stock-level 401(k) ownership, prompting investors with a preference for this characteristic to bid up their prices even further than in the absence of such preferences. Conversely, stocks facing outflows due to the shift from active to passive funds experience a decline in their stock-level 401(k) ownership, resulting in relatively lower counterfactual returns when funds prioritize this characteristic.[36]

This analysis underscores the potential scale of the transition from active to passive (captured by $NEF^{401k}(n)$ (%)) and highlights the importance of the feedback effect due to stock-level 401(k) ownership, which amplifies the impact. Figure D.2 shows similar results when all investors, not just funds, display preferences for 401(k) ownership. Since

[36]The distribution of counterfactual returns is nearly linear for stocks experiencing inflows, showing consistently higher appreciation compared to the benchmark scenario where preferences for 401(k) ownership are irrelevant. In contrast, the relationship becomes concave for stocks with outflows, exhibiting larger depreciation as benchmark returns increase, illustrating the asymmetric nature of the feedback effect.

the findings are consistent across both scenarios, we focus our discussion on the case where only funds exhibit a preference for stock-level 401(k) ownership to streamline the analysis and highlight the key dynamics.

Table 7 presents the counterfactual returns, both equally- and value-weighted, for quintile portfolios sorted by four characteristics: $NEF^{401k}(n)$ (%), size, and book-to-market. These returns are presented under two scenarios: when investors exhibit a preference for 401(k) stock ownership (even columns) and when they do not (odd columns).[37] Focusing on the NEF variable (first panel) and the benchmark case where there is no preference for 401(k) ownership, we observe that portfolios with net expected outflows (Low NEF) experience an average decline of 10%, while those with net expected inflows (High NEF) gain 21%. These results reflect the direct impact of flow-induced price pressure. More importantly, when funds exhibit a preference for 401(k) stock ownership (columns (2) and (4)), the repricing effects intensify significantly: portfolios with net expected outflows decline by 20%, while those with net expected inflows appreciate by 35%. This is consistent with the fact that stocks with outflows (Low NEF) tend to have higher investor-weighted $\beta_i^{IO^{401k}}$ (about 1.8 vs. 0.8) compared to stocks with inflows (High NEF). Thus, stocks that experience outflows, coupled with a decrease in stock-level 401(k) ownership, face a representative owner who exhibits a high coefficient on stock-level 401(k) ownership. As a result, these stocks are subject to significant negative price pressure from institutional investors. Interestingly, we can calculate the average NEF of our NEF-sorted portfolios and estimate the flow multiplier, following Gabaix and Koijen (2023). The NEF values for our value-weighted portfolios are -2.4%, -0.4%, and 0.9%, respectively, implying a flow multiplier of approximately 12-14 (e.g., 12.68%/0.9% and -29%/-2.4%). This estimate exceeds the multiplier value of 5 reported by Gabaix and Koijen (2023), suggesting that markets may be even more inelastic than previously thought.

Moving to the size sort (second panel), we observe that mid-cap stocks generally depreciate, while large-cap stocks appreciate as 401(k) plan assets shift from active to passive funds.[38] This pattern arises because large-cap stocks typically have positive $NEF^{401k}(n)$

[37] Figure D.2 illustrates the scenario where all investors, not just active funds, display preferences for stock-level 401(k) ownership. The conclusions remain nearly identical.

[38] Small-cap stocks show a modest positive effect, with returns ranging from 2% to 7%, though only

(%) values (e.g., 0.13%), while smaller stocks have negative values (e.g., −0.63%). Notably, the depreciation of mid-cap stocks intensifies when investors exhibit a preference for stock-level 401(k) ownership, amplifying the transition's impact on this group of stocks. This aligns with mid-cap stocks exhibiting a higher investor-weighted $\beta_i^{IO^{401k}}$ (about 1.5) compared to smaller or larger stocks (about 1.1).

The third panel presents the results of counterfactual returns for portfolios sorted by book-to-market ratios. Returns are negative for growth stocks and positive for value stocks, with these effects becoming significantly more pronounced under the scenario where investors care about 401(k) ownership, accompanied by a rise in statistical significance. The depreciation of growth stocks aligns with the observed decline in the value premium over the past 20 years, as active funds have increasingly tilted toward growth stocks.

4.3.1 Dissecting the Repricing

The feedback effect of stock-level 401(k) ownership on repricing is expected to be particularly pronounced when a firm's largest investors exhibit a strong preference for this characteristic, as this amplifies the institutional price pressure the firm is likely to experience. To test this, for each stock n we construct the investor-weighted stock-level 401(k) ownership coefficient:

$$\sum_i w_i(n)\beta_i^{IO^{401k}}$$

In the top panel of Figure 7, we plot the counterfactual returns (on the y-axis) when funds display a preference for stock-level 401(k) ownership as a function of the investor-weighted stock-level 401(k) ownership coefficient. Stocks overweighted by active funds offered by pension plans (according to equation (16)) are shaded in different colors based on $NEF^{401k}(n)$ (%), with black crosses indicating the stocks experiencing the largest net expected outflows.

The figure demonstrates that for stocks with significant net expected outflows (e.g., triangles and crosses), a stronger preference for stock-level 401(k) ownership by the stock's

marginally significant.

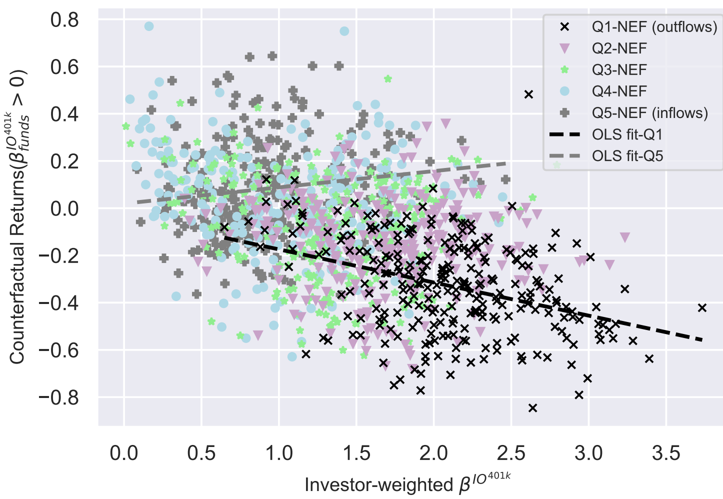
largest owners corresponds to lower counterfactual prices. In other words, stocks that are overweighted by active funds experience depreciation following outflows toward passive investments, with this effect becoming more pronounced as aggregate investors' preference for stock-level 401(k) ownership increases. Conversely, the relationship reverses for stocks underweighted by active funds, which have significant net expected inflows (e.g., circles and diamonds). In this case, if a firm's representative owner places a high value on stock-level 401(k) ownership, the firm experiences higher returns, likely driven by positive changes in its stock-level 401(k) ownership.[39]

The bottom panel of Figure 7 plots the counterfactual return, when funds exhibit a preference for stock-level 401(k) ownership ($\beta^{IO^{401k}}_{funds} > 0$), against the *change* in stock level 401(k) ownership on the x-axis.[40] This change is computed as the difference between the new counterfactual 401(k) ownership—evaluated at equilibrium prices and weights—and the observed 401(k) ownership before the counterfactual reallocation. Each dot is colored based on the investor-weighted beta of stock-level 401(k) ownership, with blacks crosses representing stocks with higher loading on this characteristic. As expected, we observe that a larger positive change in 401(k) ownership corresponds to higher returns. This indicates that stocks underweighted by active funds experience positive price appreciation not only due to inflows but also because the value of stock-level 401(k) ownership increases as a result of the shift to passive funds—a factor that becomes significant when investors care about this characteristic. Additionally, several stocks exhibit a negative change in 401(k) ownership, ranging from -0.06 to -0.02. Notably, stocks with a significant decrease in 401(k) ownership (on the far left of the x-axis) have often investors with a strong preference for the 401(k) ownership characteristic. In contrast, stocks with a large increase in 401(k) ownership (far right of the x-axis) tend to be owned by investors with a weaker preference for 401(k) ownership. This observation aligns with intuition: investors who strongly favor stock-level 401(k) ownership are more likely to choose stocks that are

[39]The two stock groups (overweighted vs. underweighted by active funds) have distinct slopes, as shown in the graph.

[40]When investors do not consider stock-level 401(k) ownership ($\beta^{IO^{401k}} = 0$ for all investors), the relationship between counterfactual returns and stock-level 401(k) ownership is a flat line. As a result, changes in stock-level 401(k) ownership have no effect on counterfactual prices. See Appendix Figure D.3.

overweighted by active funds in pension plans, as those stocks will have larger stock-level 401(k) ownership. When active funds are shut down in our counterfactual, these stocks are those that will experience a large reduction in stock-level 401(k) ownership levels.



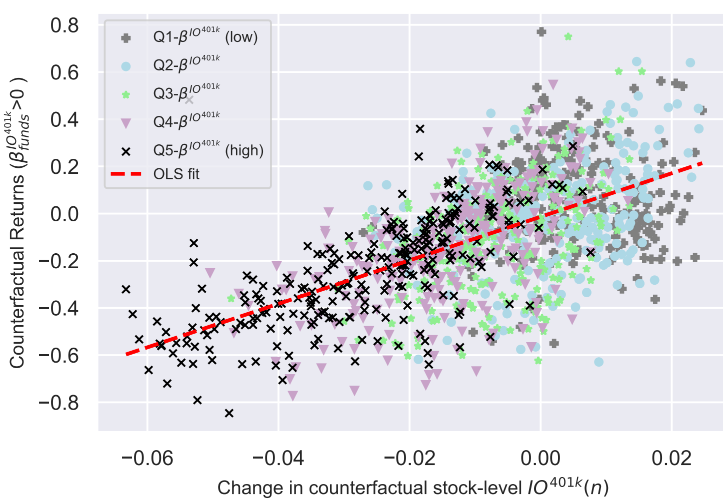


Figure 7: Counterfactual returns as a function of $\beta^{IO^{401(k)}}$ and stock-level 401(k) ownership. The top panel of this figure presents counterfactual returns (y-axis) plotted against the investor-weighted stock-level 401(k) ownership preference coefficient. Dots are shaded based on $NEF^{401k}(n)$ (%), with black dots indicating stocks experiencing the largest net expected outflows. The bottom panel displays counterfactual returns against changes in stock-level 401(k) ownership (x-axis). Each dot is colored according to the investor-weighted beta of stock-level 401(k) ownership, with black dots representing stocks with the highest exposure to this characteristic. The counterfactual return is defined as the difference between the log of the counterfactual price—when mutual funds consider stock-level 401(k) ownership ($\beta^{IO^{401k}}_{funds} > 0$)—and the log of the observed price before the reallocation of AUM from active to passive funds.

Table 8 further quantifies the importance of stock-level 401(k) ownership in shaping equity prices. Specifically, we estimate a panel regression of stock valuation variables on standard characteristics from our demand system, augmented with stock-level 401(k) ownership:

$$y_t(n) = \alpha + \boldsymbol{\beta}_1' \mathbf{X}_t(n) + \beta_2 IO^{401k}(n) + \eta_t + \gamma(n) + \epsilon_t(n) \tag{17}$$

where $y_t(n)$ is either the log market-to-book $mb_t(n)$ or the counterfactual return of stock n, η_t are time (year) fixed effects, and $\gamma(n)$ are firm-fixed effects. Column (1) uses the stock-level 401(k) ownership, while the other specifications (columns 2-5) use changes in counterfactual stock-level 401(k) ownership. All regressors are standardized. The first column reports the regression coefficients for the baseline specification over the full sample. The six characteristics in the baseline specification explain most of the cross-sectional variation in market-to-book equity, with an adjusted R^2 of 82%. Notably, market-to-book equity increases with stock-level 401(k) ownership: a one standard deviation increase in stock-level 401(k) ownership corresponds to a 9.7% increase in market-to-book equity. Additionally, the negative coefficient on log book equity implies that smaller firms have higher market-to-book ratios, with our loading on book equity closely aligning with results from Koijen et al. (2024). To assess the impact on equity prices, the second and third columns present changes in the regression coefficients resulting from reallocating assets from active to passive funds within pension plans. Specifically, the second column is based on regressing the counterfactual returns under the first scenario, e.g., $\beta^{IO^{401k}}_{funds} > 0$, on the change between the *counterfactual* and actual stock level 401(k) ownership, e.g., $\Delta IO^{401k}_t(n)$. Only the coefficient on changes in stock-level 401(k) ownership is significant, explaining nearly 40% of the counterfactual returns.[41]

Importantly, we can further decompose the difference between the counterfactual and actual stock-level 401(k) ownership into changes due to flows at the prior price and

[41] The coefficients on other exogenous characteristics remain relatively stable in the counterfactual, suggesting approximately linear effects, as detailed in Sections 2 and 7 of Koijen et al. (2024).

changes due to repricing:

$$
\begin{aligned}
\Delta IO_t^{401k}(n) &= \frac{\sum_{j\in\mathcal{P}}\left(\sum_{p=1}^{M} AUM_{p,j,t}^{NEW}\right)\times w_{j,t}^{CF}(n)}{P_t^{CF}(n)} - \frac{\sum_{j\in\{\mathcal{A},\mathcal{P}\}}\left(\sum_{p=1}^{M} AUM_{p,j,t}\right)\times w_{j,t}(n)}{P_t(n)} \\
&= \underbrace{\frac{\sum_{j\in\mathcal{P}}\left(\sum_{p=1}^{M} AUM_{p,j,t}^{NEW}\right)\times w_{j,t}^{CF}(n)}{P_t(n)} - \frac{\sum_{j\in\{\mathcal{A},\mathcal{P}\}}\left(\sum_{p=1}^{M} AUM_{p,j,t}\right)\times w_{j,t}(n)}{P_t(n)}}_{\text{changes in 401k due to flow at the old price}} + \\
&\quad \underbrace{\sum_{j\in\mathcal{P}}\left(\sum_{p=1}^{M} AUM_{p,j,t}^{NEW}\right)\times w_{j,t}^{CF}(n)\times\left(\frac{1}{P_t^{CF}(n)}-\frac{1}{P_t(n)}\right)}_{\text{changes in 401k due to the new price}} \qquad (18)
\end{aligned}
$$

The last two columns of Table 8 demonstrate that both the direct effect of reallocating flows and the repricing effect are significant contributors to variations in stock-level 401(k) ownership, which in turn drive changes in counterfactual returns.

4.3.2 Impact on price informativeness

The shift from active to passive investment management may affect price informativeness, as discussed by Koijen et al. (2024). Following Bai et al. (2016), we measure price informativeness based on a cross-sectional regression of future profitability on the ratio of market equity to book assets:

$$
\frac{E_{t+3}(n)}{A_t(n)} = \alpha + \pi \log\left(\frac{P_t(n)}{A_t(n)}\right) + \rho\frac{E_t(n)}{A_t(n)} + \nu_t(n). \qquad (19)
$$

The coefficient π measures price informativeness, with higher values indicating more informative prices.

For 2018, the standardized coefficient is 0.013, meaning a one standard deviation change in the market-to-book ratio predicts a 1.3 percentage point change in profitability.[42]

We re-estimate the cross-sectional regression (19), replacing actual market equity in 2018 with the counterfactual market equity following a shift from active to passive strategies within 401(k) plans.

[42] We focus on 2018 to exclude the COVID period but find similar results for other years.

This sharp decline in price informativeness contrasts with Koijen et al. (2024), underscoring the importance of counterfactual design and the source of counterfactual flows. Our reallocation is driven by 401(k) assets shifting from active to passive funds, whereas Koijen et al. (2024) compute counterfactual equity prices for Q4 2016, assuming institutional wealth distribution remained at Q4 2007 levels.[43]

[43]This difference across counterfactuals is also reflected in the repricing effect, with our estimate (30%) nearly double that of Koijen et al. (2024).

5 Conclusion

This paper examines how the reallocation from active to passive strategies within 401(k) plans affects asset prices. Using novel granular data on 401(k) fund offerings, merged with mutual fund holdings, we first show that stock-level 401(k) ownership is a key determinant of demand for individual stocks by mutual funds, alongside firm size. A one standard deviation increase in 401(k) stock ownership leads to a 54% increase in demand from active mutual funds (t-stat: 5.65), translating to a change in portfolio weight from 3.0% to 4.6%. This effect is distinct from other forms of institutional ownership, highlighting the unique informational content of 401(k) ownership for fund managers.

Building on this evidence, we develop a novel counterfactual demand-based asset pricing methodology to assess the equilibrium price impact of the shift of 401(k) plans toward passive funds. Unlike traditional models that treat firm characteristics as exogenous, our framework captures the feedback loop between 401(k) ownership and investors' investment decisions, allowing both stock prices and 401(k) stock ownership to evolve endogenously.

Our counterfactual analysis reveals that the transition to passive funds has substantial effects on equity prices, with stocks previously overweighted by active funds experiencing significant declines due to direct outflows and reduced stock-level 401(k) ownership (indirect effect). In a benchmark scenario where funds do not consider stock-level 401(k) ownership, stocks facing outflows decline by 10%, while those receiving inflows gain 20%. However, when investors exhibit preferences for stock-level 401(k) ownership, price effects are amplified: stocks with outflows decline by 20%, while those with inflows rise by 35%. Additionally, we estimate a price-flow multiplier of 12–14, exceeding existing estimates of 5 (Gabaix and Koijen, 2023), indicating that equity markets may be even more inelastic than previously thought if drastic changes in the allocation of a subset of investors (defined contribution plans) take place. Finally, our results highlight heterogeneous effects across stock characteristics—large-cap and value stocks tend to benefit from this shift to passive within 401(k) plans, while mid-cap stocks experience strong declines, particularly when investors favor stocks with high 401(k) ownership.

Taken together, our findings demonstrate that pension assets not only directly impact stock prices but also, a key finding of our paper, influence investors' stock demand. These insights are critical in understanding the broader implications of the ongoing transition from active to passive investing, particularly in the context of retirement savings. By introducing a quantitative framework to assess these dynamics, our study contributes to the growing literature on demand-based asset pricing by emphasizing the evolving role of pension assets in financial markets.

References

Andrews, Isaiah, James H Stock, and Liyang Sun (2019) "Weak Instruments in Instrumental Variables Regression: Theory and Practice," *Annual Review of Economics*, Vol. 11, pp. 727–753.

Backus, Matthew, Christopher Conlon, and Michael Sinkinson (2021) "Common Ownership in America: 1980–2017," *American Economic Journal: Microeconomics*, Vol. 13, pp. 273–308, 8.

Ben-David, Itzhak, Francesco Franzoni, Rabih Moussawi, and John Sedunov (2021) "The Granular Nature of Large Institutional Investors," *Management Science*, Vol. 67, pp. 6629–6659.

Ben-David, Itzhak, Jiacui Li, Andrea Rossi, and Yang Song (2023) "Discontinued Positive Feedback Trading and the Decline of Return Predictability," *Journal of Financial and Quantitative Analysis*.

Bhamra, Harjoat S. and Raman Uppal (2019) "Does Household Finance Matter? Small Financial Errors with Large Social Costs," *American Economic Review*, Vol. 109, No. 3, p. 1116–1154.

Bretscher, Lorenzo, Lukas Schmid, Ishita Sen, and Varun Sharma (2021) "Institutional Corporate Bond Pricing," *Working Paper*.

Bushee, Brian J (1998) "The Influence of Institutional Investors on Myopic R & D Investment Behavior," *The Accounting Review*, Vol. 73, pp. 305–333.

Chau, Yuet, Karamfil Todorov, and Eyub Yegen (2025) "ETFs as a disciplinary device," *Working Paper*.

Chen, Hsiu-Lang, Narasimhan Jegadeesh, and Russ Wermers (2000) "The Value of Active Mutual Fund Management: An Examination of the Stockholdings and Trades of Fund Managers," *The Journal of Financial and Quantitative Analysis*, Vol. 35, pp. 343–368.

Christoffersen, Susan E K and Mikhail Simutin (2017) "On the Demand for High-Beta Stocks: Evidence from Mutual Funds," *The Review of Financial Studies*, Vol. 30, pp. 2596–2620, 9.

Egan, Mark, Alexander MacKay, and Hanbin Yang (2021) "What Drives Variation in Investor Portfolios? Evidence from Retirement Plans," *Working Paper*.

Fan, Yinghua, Guanhao Feng, Scott Au, and Sayad Baronyan (2022) "Estimating Asymmetric Price Impact," *Working Paper*.

Gabaix, Xavier and Ralph S J Koijen (2023) "In Search of the Origins of Financial Fluctuations:The Inelastic Markets Hypothesis," *Working Paper*.

Gabaix, Xavier and Ralph S. J. Koijen (2024) "Granular Instrumental Variables," *Journal of Political Economy*, Vol. 132, No. 7, pp. 2274–2303.

Gabaix, Xavier, Ralph S J Koijen, Federico Mainardi, Sangmin Oh, and Motohiro Yogo (2022) "Asset Demand of U.S. Households," *Working Paper*.

Haddad, Valentin, Paul Huebner, and Erik Loualiche (2025) "How Competitive is the Stock Market? Theory, Evidence from Portfolios, and Implications for the Rise of Passive Investing," *American Economic Review, forthcoming*.

Koijen, Ralph S J and Motohiro Yogo (2019) "A Demand System Approach to Asset Pricing," *Journal of Political Economy*, Vol. 127, pp. 1475 – 1515.

Koijen, Ralph S J, Robert Richmond, and Motohiro Yogo (2024) "Which Investors Matter for Global Equity Valuations and Expected Returns?" *The Review of Economic Studies*, Vol. 91, pp. 2387–2424, 7.

Noh, Don and Sangmin Oh (2020) "Measuring institutional pressure for greenness: A demand system approach," *Working Paper*.

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*, Vol. 71, pp. 1779–1812.

Pástor, Luboš, Robert F Stambaugh, and Lucian A Taylor (2015) "Scale and skill in active management," *Journal of Financial Economics*, Vol. 116, pp. 23–45, 4.

Schwarz, Christopher G and Mark E Potter (2016) "Revisiting Mutual Fund Portfolio Disclosure," *Review of Financial Studies*, Vol. 29, pp. 3519–3544, 9.

Sialm, Clemens and Laura Starks (2012) "Mutual Fund Tax Clienteles," *Journal of Finance*, Vol. 67, pp. 1397–1422, 8.

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

Stock, James H and Motohiro Yogo (2005) *Testing for Weak Instruments in Linear IV Regression*, pp. 80–108: Cambridge University Press.

Wermers, Russ (2000) "Mutual Fund Performance: An Empirical Decomposition into Stock-Picking Talent, Style, Transactions Costs, and Expenses," *The Journal of Finance*, Vol. 55, pp. 1655–1695.

Panel A: Distribution of fund-level IO^{401k} and stock-level $IO^{401k}(n)$				
	25th	Median	Mean	75th
IO^{401k} (all funds)	0.31%	2.72%	8.05%	10.46%
IO^{401k} (index funds)	0.85%	3.99%	9.34%	14.53%
IO^{401k} (active funds)	0.28%	2.60%	7.91%	10.04%
IO^{401k} (active MFs)	0.86%	3.87%	9.25%	12.59%
IO^{401k} (active ETFs)	0.02%	0.05%	0.89%	0.21%
Persistence on fund allocation	0.27	0.59	0.54	0.82
$IO^{401k}(n)$	1.72%	2.79%	3.09%	4.19%

Panel B: Total \$ allocation across type of funds				
	2008	2012	2016	2020
Total assets of 401(k) plans ($ bn)	835.19	2,605.92	3,851.36	5,927.89
Allocation in US equity index funds ($ bn)	32.66	130.06	269.40	380.84
Allocation in US equity active MFs ($ bn)	117.31	377.24	500.81	628.60
Allocation in US equity active ETFs ($ bn)	1.93	8.24	19.92	32.43

Table 1: Summary statistics. This table reports summary statistics of the cross-sectional distribution of 401(k) plans characteristics. IO^{401k} indicates the fraction of a fund assets collectively owned by 401(k) plans. The first row ("all funds") considers the universe of all US equity (active and index) funds. Index funds comprise mutual funds classified according to the Morningstar variables as "index funds" and "enhanced index", and ETFs with the S&P500 index as benchmark. The second and third rows only include US equity active mutual funds and ETFs, respectively. The fourth and fifth rows include the set of index (MFs and ETFs) and active (MFs and ETFs) funds, respectively. $IO^{401k}(n)$ represents the 401(k) plans ownership of stock n. Persistence on fund allocation is the AR(1) coefficient on the fraction of 401(k) plan assets invested in a specific fund. Panel B reports the total assets managed by 401(k) plans over time. The last three rows report the allocation of 401(k) plans into US equity index funds (both index MFs and ETFs), US equity active MFs, and US equity active ETFs, respectively.

Panel A: Funds owned by 401(k) pension plans												
	All Funds				*Mutual Funds*				*ETFs*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\widehat{IO}_{-i,t}^{401k}(n)$	0.32***	0.32***	0.30***	0.27***	0.54***	0.52***	0.50***	0.47***	0.12**	0.13**	0.12**	0.08
	(4.71)	(4.88)	(4.63)	(3.94)	(5.65)	(5.46)	(5.33)	(5.16)	(2.59)	(2.59)	(2.48)	(1.61)
Log market-to-book	0.85***	0.85***	0.86***	0.84***	0.34***	0.35***	0.35***	0.32***	0.96***	0.96***	0.96***	0.94***
	(10.65)	(10.61)	(10.54)	(9.91)	(4.34)	(4.53)	(4.61)	(4.31)	(17.23)	(17.17)	(17.06)	(16.93)
Log book equity	0.98***	0.98***	0.98***	0.97***	0.64***	0.65***	0.65***	0.63***	1.01***	1.01***	1.01***	1.0***
	(24.41)	(24.09)	(23.8)	(24.35)	(12.4)	(12.78)	(12.7)	(12.63)	(43.46)	(42.29)	(41.91)	(42.8)
Operating profitability	0.13***	0.13***	0.13***	0.12***	0.11***	0.12***	0.12***	0.12***	0.12***	0.12***	0.12***	0.12***
	(7.0)	(6.93)	(6.76)	(6.34)	(5.62)	(5.83)	(5.9)	(5.85)	(5.59)	(5.62)	(5.52)	(5.33)
Beta	-0.1***	-0.1***	-0.1***	-0.1***	-0.07***	-0.07***	-0.06***	-0.06***	-0.08***	-0.08***	-0.08***	-0.08***
	(-5.22)	(-5.25)	(-4.99)	(-5.46)	(-5.83)	(-5.84)	(-5.17)	(-5.73)	(-4.36)	(-4.39)	(-4.21)	(-4.45)
Investment	-0.08***	-0.08***	-0.08***	-0.08***	-0.05***	-0.05***	-0.05***	-0.05***	-0.07***	-0.07***	-0.07***	-0.07***
	(-8.16)	(-8.18)	(-8.04)	(-8.19)	(-3.65)	(-3.74)	(-3.68)	(-3.66)	(-5.28)	(-5.31)	(-5.29)	(-5.24)
Dividend-to-book	0.08***	0.08***	0.08***	0.1***	-0.01	-0.0	0.0	0.01	0.09***	0.09***	0.1***	0.11***
	(6.58)	(6.61)	(6.7)	(7.47)	(-0.44)	(-0.09)	(0.28)	(0.33)	(6.12)	(6.11)	(6.12)	(7.06)
Top10 ownership		0.01	-0.0			0.07***	0.04***			-0.01	-0.02	
		(0.47)	(-0.05)			(7.12)	(3.93)			(-0.44)	(-0.87)	
Mutual Fund ownership			0.05***				0.11***				0.04***	
			(4.21)				(4.84)				(4.25)	
DED				0.01				0.01				0.0
				(0.34)				(0.31)				(0.01)
QIX				0.1***				0.12***				0.09***
				(7.51)				(5.69)				(7.16)
TRA				0.04*				-0.0				0.02
				(1.83)				(-0.04)				(0.89)

Table 2: Demand system estimation - Stock level $IO_t^{401k}(n)$ (continued on the next page).

Panel B: Funds not managing 401(k) assets												
	All Funds				*Mutual Funds*				*ETFs*			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$\widehat{IO}^{401k}_{-i,t}(n)$	0.45***	0.46***	0.45***	0.41***	0.46***	0.47***	0.46***	0.42***	0.35**	0.36**	0.36**	0.32**
	(4.74)	(4.87)	(4.83)	(4.54)	(4.47)	(4.65)	(4.57)	(4.31)	(2.57)	(2.62)	(2.63)	(2.36)
Log market-to-book	0.68***	0.68***	0.68***	0.68***	0.53***	0.52***	0.52***	0.52***	0.84***	0.84***	0.84***	0.84***
	(12.25)	(12.16)	(12.11)	(12.22)	(7.97)	(7.92)	(7.92)	(8.02)	(10.23)	(10.16)	(10.16)	(10.12)
Log book equity	0.78***	0.77***	0.77***	0.78***	0.68***	0.67***	0.67***	0.68***	0.88***	0.88***	0.88***	0.88***
	(20.3)	(19.9)	(19.79)	(20.23)	(16.48)	(16.43)	(16.43)	(16.6)	(17.08)	(16.58)	(16.58)	(17.01)
Operating profitability	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	(0.8)	(0.83)	(0.83)	(0.77)	(0.56)	(0.58)	(0.58)	(0.53)	(1.02)	(1.06)	(1.06)	(1.01)
Beta	-0.12***	-0.12***	-0.12***	-0.12***	-0.13***	-0.13***	-0.12***	-0.13***	-0.11***	-0.11***	-0.11***	-0.11***
	(-6.14)	(-6.34)	(-6.27)	(-6.25)	(-5.82)	(-6.07)	(-5.91)	(-5.96)	(-4.9)	(-4.95)	(-4.96)	(-5.0)
Investment	0.0	0.0	0.0	0.0	-0.0	-0.0	-0.0	-0.0	0.0	0.0	0.0	0.0
	(0.05)	(0.2)	(0.19)	(0.09)	(-0.41)	(-0.3)	(-0.29)	(-0.35)	(0.4)	(0.54)	(0.54)	(0.39)
Dividend-to-book	0.03***	0.03***	0.03***	0.03***	0.02***	0.02***	0.02***	0.03***	0.04***	0.04***	0.04***	0.04***
	(5.18)	(5.23)	(5.2)	(5.17)	(5.64)	(5.76)	(5.68)	(5.67)	(4.14)	(4.14)	(4.14)	(4.15)
Top10 ownership		-0.04*	-0.05**			-0.04	-0.05*			-0.05**	-0.05**	
		(-1.98)	(-2.21)			(-1.47)	(-1.83)			(-2.23)	(-2.22)	
Mutual Fund ownership			0.03**				0.05***				-0.0	
			(2.9)				(4.2)				(-0.19)	
DED				-0.03**				-0.04				-0.02
				(-2.19)				(-1.68)				(-1.39)
QIX				0.05***				0.06***				0.04***
				(4.14)				(3.57)				(3.03)
TRA				0.03				0.02				0.04
				(1.43)				(0.89)				(1.35)

Table 2: Demand system estimation - Stock level $IO^{401k}_t(n)$ (continued). This table reports estimates of the panel regression for funds not managing 401(k) assets.

$$\log\left(\frac{w_{i,t}(n)}{w_{i,t}(0)}\right) = b_0 + \beta_0 \widehat{mb}_{-i,t}(n) + \boldsymbol{\beta_1'}\mathbf{X}_t(n) + \beta^{IO^{401k}} \widehat{IO}^{401k}_{-i,t}(n) + \alpha_{i,t} + \tilde{u}_{i,t}(n)$$

where $\widehat{mb}_t(n)$ is the instrumented market-to-book equity, and $\mathbf{X}_t(n)$ includes the same variables as in Koijen and Yogo (2019), i.e., log book equity, operating profitability, stock market beta, investment, and dividend-to-book ratio. $\widehat{IO}^{401k}_{-i,t}(n)$ is the instrumented 401K plans ownership of stock n (excluding the effect through investor i), and $\alpha_{i,t}$ are fund-by-time fixed effects. The funds in the regressions are AUM-weighted. Panel A reports results for all funds owned by 401(k) plans, while Panel B only includes funds not managing 401(k) assets (hence, $IO^{401k}_{-i,t}(n) \equiv IO^{401k}_t(n)$). Columns (1)-(3) report results using all funds (including index mutual funds and ETFs), columns (4)-(6) using only active mutual funds, and columns (7)-(9) using only active ETFs. Variables are standardized to have unit-standard deviation. Standard errors are triple clustered by fund, time and stock. The sample period is from 2007 to 2020. *p<0.1; **p<0.05; ***p<0.01.

Panel A: Stock-level 401(k) ownership (mutual funds)			
	Coefficient	s.e.	t-stat
$IO^{401k}_{-i,t}(n)$	0.250***	0.010	25.490
Log market-to-book	0.132***	0.034	3.890
Log book equity	0.590***	0.025	23.750
Operating profitability	0.067***	0.007	9.280
Beta	-0.124***	0.008	-14.720
Investment	-0.097***	0.007	-13.800
Dividend-to-book	0.014*	0.008	1.800

Panel B: Fund-level 401(k) ownership (mutual funds)			
	Coefficient	s.e.	t-stat
$IO^{401k}_{i,t}$	0.048*	0.027	1.770
Log market-to-book	0.199***	0.036	5.570
Log book equity	0.530***	0.027	19.780
Operating profitability	0.051***	0.007	6.980
Beta	-0.131***	0.009	-15.020
Investment	-0.094***	0.007	-13.340
Dividend-to-book	-0.037***	0.008	-4.730

Table 3: Demand system estimation - GMM with stock- and fund-level 401(k) ownership. This table reports GMM estimates of the regression

$$\frac{w_{i,t}(n)}{w_{i,t}(0)} = \exp\left\{b_0 + \beta_{0,i}\widehat{mb}_t(n) + \boldsymbol{\beta_1'}\mathbf{X}_t(n) + \beta_i^{IO^{401k}} IO_t^{401k} + \alpha_t\right\} \tilde{u}_{i,t}(n)$$

where $\widehat{mb}_t(n)$ is the instrumented log market equity-to-book equity, and $\mathbf{X}_t(n)$ includes the same variables as in Koijen and Yogo (2019), i.e., log book equity, operating profitability, stock market beta, investment, and dividend-to-book ratio. IO_t^{401k} indicates either the 401(k) plans ownership of the individual stock n excluding the effect through fund i (Panel A), or the 401(k) plans ownership of fund i (Panel B). We report results using only active mutual funds. The estimation includes observations of mutual funds with zero stock-holdings but still in the investment universe, and observations are AUM-weighted. Variables are standardized. Standard errors are double clustered by fund and time. The sample period is from 2007 to 2020. *p<0.1; **p<0.05; ***p<0.01.

Panel A: Characteristics of the Matched Sample		
	Treated Group	Control Group
Number of Stocks	8,963	8,963
Market Capitalization ($bn)	11.0	12.2
Book-to-market Ratio	0.51	0.56
Beta	1.15	1.01

Panel B: Panel Regressions			
	(1)	(2)	(3)
Treated dummy$_t$	0.050***	0.042***	0.032***
	(3.352)	(3.200)	(3.248)
Size$_{t-1}$		-0.027***	-0.027***
		(-2.866)	(-3.343)
Book-to-market$_{t-1}$			-0.056
			(-1.629)
Beta$_{t-1}$			0.024
			(0.627)
Momentum$_{t-1}$			0.018
			(0.447)
Year FEs	Yes	Yes	Yes
No. Observations	17,398	17,398	17,398

Table 4: Matching stocks: impact of 401(k) ownership. Panel A reports the average stock characteristics of the stocks in the treatment and control groups. We rank candidate control stocks based on the absolute difference in market capitalization from the treated stock, generating a market cap rank (where rank = 1 is the closest match). We apply the same ranking methodology to the book-to-market ratio. The control stock with the lowest combined rank across both criteria is selected as the match for each treated stock annually. Panel B reports results from of regression on the matched sample. After matching stocks as described in Section 4.1.1, we estimate a panel regression of annual returns of the matched pair of stocks on a treated dummy variable and various stock-level controls. Controls include lagged size, book-to-market, beta, and momentum. Standard errors are double clustered by stock and time.

	ret_t			ret_{t+1}			$ret_{t+1:t+3}$		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
401(k) dummy - Large Δ holdings	0.119***	0.113***	0.116***	-0.041***	-0.029***	-0.019***	-0.119***	-0.095***	-0.066***
	(4.015)	(3.849)	(4.563)	(-2.682)	(-2.650)	(-2.884)	(-4.074)	(-8.293)	(-4.270)
401(k) dummy - Small Δ holdings		-0.024	-0.033		0.049	0.039		0.098	0.092
		(-0.641)	(-0.886)		(1.124)	(0.964)		(1.128)	(1.124)
Top 10 investors dummy - Large Δ holdings			0.002			-0.018			-0.095***
			(0.100)			(-1.027)			(-2.655)
Top 10 investors dummy - Small Δ holdings			0.054***			0.068**			0.061
			(3.011)			(2.531)			(1.552)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 5: Stock-level 401(k) trading and returns. This table reports estimates of regressions of stock returns on changes in 401(k) plans and top 10 institutions' holdings. The dependent variables are contemporaneous returns (columns (1)-(3)), next year (t+1) returns (column (4)-(6)), and cumulative t+1:t+3 returns (column (7)-(9)). Controls include log market equity and time fixed effects. Standard errors are double clustered by stock and time.

	(1)	(2)	(3)	(4)
$\widehat{\text{Demand}_t^{401(k),VW}}(n)$	0.481***	0.401***	0.472***	0.374***
	(7.070)	(10.070)	(8.510)	(9.330)
Size_{t-1}			-0.026***	-0.291***
			(-4.390)	(-4.920)
Beta_{t-1}			-0.003	0.011
			(-0.110)	(0.570)
$\text{Book-to-market}_{t-1}$			-0.056	0.035
			(-1.320)	(0.730)
Momentum_{t-1}			-0.027	-0.070*
			(-0.650)	(-1.950)
Stock FEs	No	Yes	No	Yes
Year FEs	Yes	Yes	Yes	Yes

Table 6: Granular instrumental variable regression. This table reports estimates of the GIV stock-level panel regression

$$r_t(n) = \beta_0 + \beta_1(n) \times \left(\widehat{\text{Demand}_t^{401(k),VW}}(n)\right) + \varepsilon_t(n)$$

The dependent variable is annual stock returns in year t. The variable of interested is 401(k) plans' demand, instrumented by GIV $\widehat{\text{Demand}_t^{401(k),VW}}(n)$. Standard errors are double clustered by stock and time.

Characteristics		Equally-weighted		Value-weighted	
		$\beta^{IO^{401k}} = 0$	$\beta^{IO^{401k}}_{funds} > 0$	$\beta^{IO^{401k}} = 0$	$\beta^{IO^{401k}}_{funds} > 0$
		Flow only	*Flow + Preference*	*Flow only*	*Flow + Performance*
		(1)	(2)	(3)	(4)
NEF					
	Low	-10.44***	-19.97***	2.24	-28.93***
		(-5.20)	(-17.94)	(0.11)	(-19.67)
	Mid	2.53	-1.09	-7.61	-15.52***
		(0.50)	(-0.67)	(-1.29)	(-8.51)
	High	21.46***	35.06***	7.48***	12.68***
		(3.98)	(5.32)	(3.10)	(3.33)
Size					
	Small	3.52	7.50*	2.36	5.21
		(1.18)	(1.82)	(0.89)	(1.59)
	Mid	-4.44**	-6.12***	-4.79***	-6.64***
		(-2.56)	(-3.53)	(-2.98)	(-3.90)
	Large	19.97***	17.53***	5.04	1.14
		(2.79)	(2.86)	(1.35)	(0.49)
Book-to-market					
	Low	0.73	-0.42	-1.09	-11.43***
		(0.29)	(-0.10)	(-0.17)	(-6.07)
	Mid	11.15	1.58	12.41*	2.44
		(1.59)	(0.32)	(1.76)	(0.49)
	High	1.65	12.81**	0.16	30.04**
		(0.59)	(2.02)	(0.04)	(2.03)

Table 7: Counterfactual portfolio returns from 401(k) asset reallocation. This table reports the equally-weighted and value-weighted average returns of portfolios sorted on $NEF^{401k}(n)$ (%) and stock characteristics. Counterfactual returns are computed at each year-end, reflecting the impact of reallocating 401(k) assets from active to passive funds. These returns account for both the pre-counterfactual and post-counterfactual market capitalizations, capturing the value change in portfolios resulting from the flow of funds and potential preferences for stock-level 401(k) ownership. Sample covers 2010-2019.

Characteristic	Actual mb (1)	Counterfactuals $\beta^{IO^{401k}}_{funds} > 0$ *cf ret* (2)	$\beta^{IO^{401k}}_{ALL} > 0$ *cf ret* (3)	$\beta^{IO^{401k}}_{funds} > 0$ *cf ret* (4)	$\beta^{IO^{401k}}_{ALL} > 0$ *cf ret* (5)
Log book equity	-0.594***	-0.029	-0.026	-0.029	-0.022
	(-14.433)	(-1.952)	(-1.916)	(-1.947)	(-1.770)
Operating profitability	0.193***	-0.004	0.001	-0.004	0.000
	(7.349)	(-0.851)	(0.239)	(-1.220)	(-0.019)
Beta	-0.044**	-0.004	-0.002	-0.004	-0.002
	(-2.846)	(-0.693)	(-0.367)	(-0.683)	(-0.274)
Investment	-0.031***	-0.001	-0.001	-0.002	-0.001
	(-7.411)	(-0.387)	(-0.297)	(-0.495)	(-0.294)
Dividend-to-book	0.040***	0.005	0.000	0.006	0.002
	(5.021)	(0.640)	(0.044)	(0.851)	(0.258)
$IO^{401k}_t(n)$	0.097***				
	(10.364)				
$\Delta IO^{401k}_t(n)$		0.156***	0.173***		
		(25.574)	(19.285)		
$\Delta IO^{401k}_t(n)$-Flow				2.708***	3.591***
				(25.073)	(19.110)
$\Delta IO^{401k}_t(n)$-Price				2.570***	3.417***
				(23.162)	(18.170)
Adjusted R^2	0.82	0.39	0.39	0.55	0.57
Observations	18,488	18,488	18,488	18,488	18,488

Table 8: Valuation regressions.This table presents the results of the regression:

$$y_t(n) = \alpha + \boldsymbol{\beta_1'}\mathbf{X}_t(n) + \beta_2 IO^{401k}(n) + \eta_t + \gamma(n) + \epsilon_t(n)$$

where $y_t(n)$ represents either the log market-to-book ratio $mb_t(n)$ (column (1)) or the counterfactual return. Column (1) regresses the actual log market-to-book equity on firm characteristics for the full 2007–2019 sample. Column (1) uses the actual stock-level 401(k) ownership, while columns (2)-(5) instead use the change in stock-level 401(k) ownership under counterfactual scenarios, where either (i) mutual fund investors prefer stocks with higher 401(k) ownership (columns 2 and 4) or (ii) all investors, except households, exhibit this preference (columns 3 and 5). All regressors have been standardized cross-sectionally by year and all specifications include year and firm fixed effects. The t-statistics, based on standard errors clustered by firm and year, are reported in parentheses.

Internet Appendix A Data Cleaning Procedure

A.1 BrightScope and Morningstar (MS)

1. We match funds held in 401(k) plans with Morningstar holdings.

2. We remove mutual funds whose portfolio weights as reported by Morningstar are different from the correct portfolio weights calculated using holdings values and total net assets, as in Pástor et al. (2015).

3. We merge fund characteristics (e.g., fund TNA) from Morningstar with the dollar allocation of 401(k) plans to funds from BrightScope. We then calculate our $IO_{i,t}^{401k}$ variable, a fund's 401(k) ownership. We drop funds where $IO_{i,t}^{401k} < 0$ or $IO_{i,t}^{401k} > 1$.

4. Our analysis focuses on equities, hence we only keep equity mutual funds having an equity ratio ≥ 0.75.

5. We merge fund holdings with firm data from CRSP and COMPUSTAT, replacing missing dividends as zero.

6. We drop fund-stock observations with missing characteristics.

7. We define the investment universe for each fund as described in the paper. We only keep funds with clearly defined investment universes (e.g., the number of stocks in the investment universe is greater than zero)

8. We drop funds holding fewer stocks than the fifth percentile in the cross-section of funds, every year (approx. 15 stocks).

9. As in Koijen and Yogo (2019), each year, we winsorize profitability, investment, and market beta at the 2.5th and 97.5th percentiles to reduce the impact of outliers. Since dividends are positive, we winsorize dividends to book equity at the 97.5th percentile. We also winsorize log(book equity) at the 2.5th and 97.5th percentiles.

10. We winsorize funds' total net assets (TNA) at the 97.5th percentile, every year, to deal with outliers.

11. In the GMM estimation, we keep zero-weight holdings, e.g., stocks in a fund's investment universe, but currently not being held by the fund. Zero-weight holdings must have non-missing characteristics.

Estimation

- In the pooled regressions, we implement 2SLS with instrumented log market-to-book, and use fund TNA as weights.
- For GMM, we include zero holdings of a stock, and use fund TNA as weights.
- As in Koijen et al. (2024), we impose the economic constraint log(MB) < 1 in all the estimations.
- The price impact analysis is based on yearly GMM estimations.

A.2 Thomson Reuters s34 Holdings

1. We use the same institutional types as in Koijen and Yogo (2019).
2. We merge the s34 holdings data with CRSP and COMPUSTAT.
3. We define the investment universe for each institution.
4. In the GMM estimation of price impact, we pool institutions into groups by type and TNA as in Koijen and Yogo (2019), include holdings with zero weights (e.g., belonging to the investor's investment universe, but not currently owned), and calculate the instrument based on these pooled groups.

A.3 Scraped Holdings from 13F Filings

1. We follow Backus et al. (2021), and use their 13F scraped holdings between 2007 and 2016.
2. We merge these holdings with CRSP and COMPUSTAT, and define the investment universe for each institution.

3. We drop institutions holdings less than 100 stocks at any given time, and pool institutions into groups by TNA as in Haddad et al. (2025). We then calculate the instrument based on these pooled groups.

4. We estimate the price impact via GMM, including holdings with zero weights (e.g., belonging to the investor's investment universe, but not currently owned).

Internet Appendix B Merging Morningstar and S34 Holdings

Following the approach of Koijen and Yogo (2019), we classify investors into four categories: institutional investors (e.g., include banks, insurance companies, and hedge funds), index mutual funds, active equity mutual funds, and households. We obtain mutual fund holdings data from Morningstar and institutional holdings data from Thomson Reuters S34 files. It is important to note that the S34 dataset aggregate holdings at the mutual fund *family* level, combining all holdings across funds within the same family. To prevent double counting, it is hence essential to merge the Morningstar and S34 holdings data. This merging process is complicated by the absence of a direct key linking the two datasets.

The initial step involves constructing a reliable fund-level mapping between Morningstar holdings and CRSP mutual funds, utilizing fund tickers, CUSIPs, and fund names. We ensure the accuracy of these matches by comparing assets and returns across the two databases, following the methodology described in Pástor et al. (2015). Ultimately, we develop a mapping that links Morningstar mutual funds (*fundid*) to CRSP mutual funds (*crsp_portno* and *crsp_cl_grp*).

Next, we map the CRSP mutual fund dataset to Thomson-Reuters mutual fund holdings S12 database using the "MFLinks" tables constructed by Wermers (2000). The MFLinks tables offer a reliable connection between CRSP mutual funds (*crsp_fundno*) and Thomson-Reuters mutual fund holdings (*fundno*) through the Wharton Financial Institution Center Number (*wficn*). Finally, we utilize the S12TYPE5 data to further link Thomson Reuters mutual fund holdings (S12) to S34 institutional holdings.

After merging the datasets, we exclude Morningstar holdings from the S34 institutional holdings. The household sector is subsequently calculated as the residual, defined as the difference between shares outstanding and the sum of the four categories of holdings. In rare cases where the sum of institutional holdings exceeds the number of shares outstanding, due to the exclusion of short positions in the S34 data or potential reporting errors, we proportionally rescale all S34 institutional holdings. This adjustment ensures that the combined shares from S34 and Morningstar mutual funds equal the total number of outstanding shares.

Internet Appendix C

C.1 Additional Plots

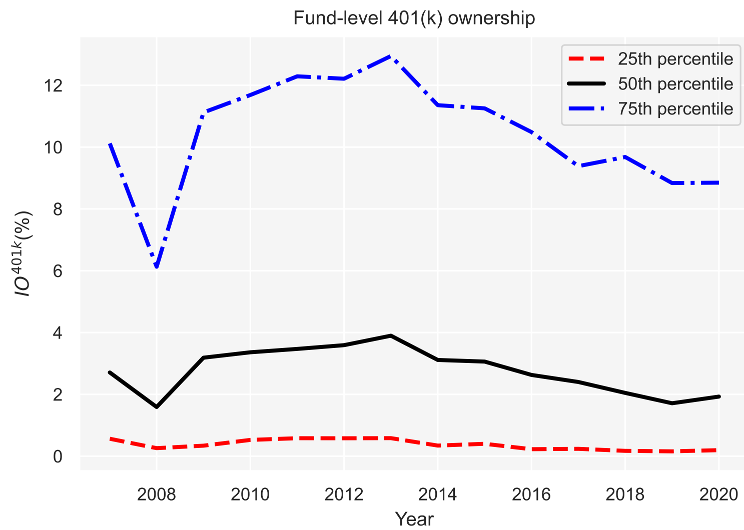


Figure C.1: Fund-level 401(k) ownership over time. This figure shows the cross-sectional distribution of fund-level 401(k) ownership over time. Annual data, from 2007 to 2020.

C.2 Coefficients on Other Characteristics

Figure C.2 shows annual estimates of the coefficients on market-to-book and the other characteristics for mutual funds (blue dotted line) and ETFs (red solid line) for the demand system that includes stock-level 401(k) ownership (see equation (6)).

To validate our estimation, we also report the coefficient estimates for index (mutual and ETF) funds (c.f., Section 2). If the estimation of our characteristics-based demand system is valid, one should recover a unit coefficient on log market equity, and zero on the other characteristics for an hypothetical index fund. Albeit the coefficient on market equity (which can be obtained from the coefficient on log market-to-book equity and log book equity) is not exactly one, we still notice that index funds are inelastic, and substantially more so than active mutual funds and ETFs. Furthermore, the coefficient of index funds on other characteristics is close to zero, the sole exception being the dividend-to-

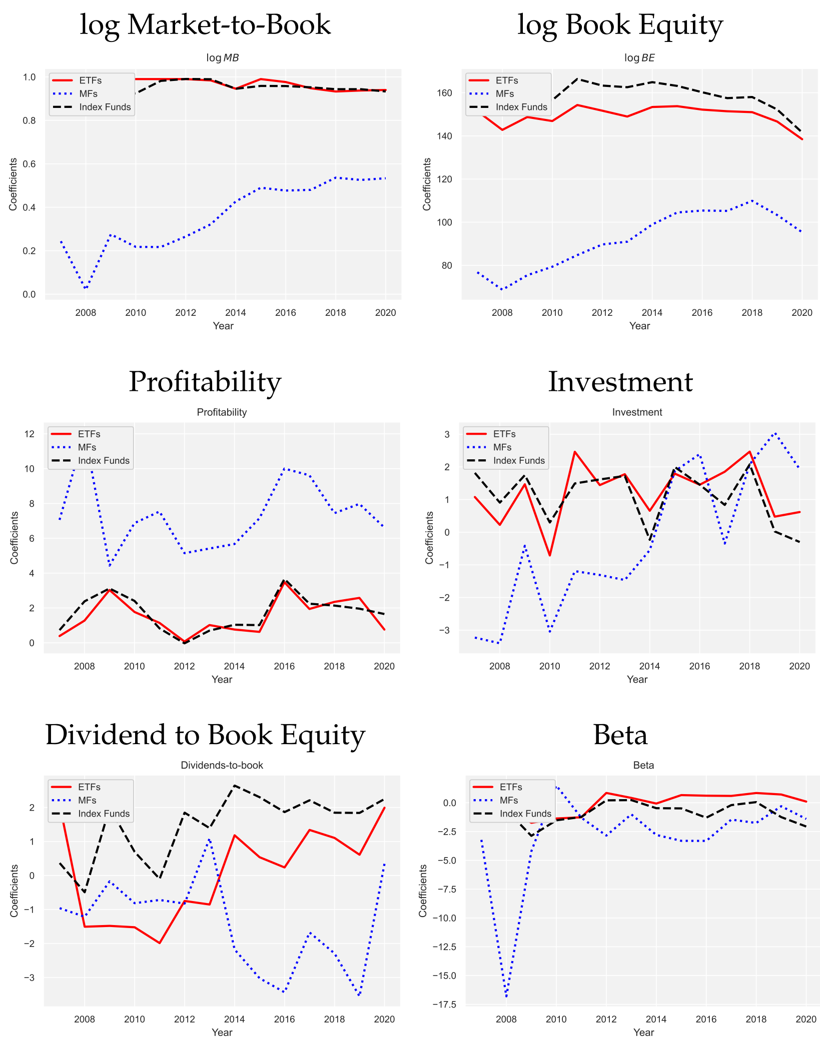


Figure C.2: Coefficients on the other characteristics - stock level. This figure shows the annual coefficients in (6), separately for mutual funds, ETFs, and index funds, estimated by pooled OLS using assets under management as weights. The regression is estimated year by year. Except for log market-to-book equity, we standardize characteristics (within each year) and multiply the coefficients by 100, so that they can be interpreted as the percentage change in demand per one standard deviation change in the characteristic. The sample period is from 2007 through 2020.

book equity. Thus we confirm the validity of our characteristics-based demand estimation and of our criteria to categorize index funds.

C.3 Robustness for Stock-Level $IO^{401k}_{i,t}(n)$ Analysis

We repeat the analysis in Table 2 without weighting observation by AUM (see Table C.1) or avoiding winsorization of fund TNA (Table C.2). Across specifications, the coefficient on $IO^{401k}_{t}(n)$ is large and statistically significant. The coefficient is also not affected by the inclusion alternative ownership variables. For example, the coefficient is 0.109 without onwership controls and 0.097 when we include the mutual fund and top-10 ownership variables (columns (4) and (6) in Panel B of Table C.1). Also, the coefficients for mutual funds and ETFs are 0.109 (t-stat=6.09) and 0.069 (t-stat=3.27), respectively, thus confirming a stronger effect for the former. For this result, however, the TNA winsorization matters. This is shown in Table C.2 where – with no winsorization – the mutual funds and ETFs coefficients get closer in magnitude.

Panel A: Funds owned by 401(k) pension plans												
	All Funds				**Mutual Funds**				**ETFs**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$IO^{401k}_{-i,t}(n)$	0.42***	0.42***	0.41***	0.38***	0.41***	0.41***	0.4***	0.37***	0.15***	0.15***	0.14***	0.11**
	(10.96)	(11.29)	(11.02)	(10.36)	(7.81)	(7.82)	(7.61)	(7.27)	(3.48)	(3.49)	(3.37)	(2.78)
Log market-to-book	0.44***	0.44***	0.44***	0.43***	0.18***	0.19***	0.19***	0.18***	0.7***	0.7***	0.7***	0.69***
	(8.59)	(8.6)	(8.53)	(8.63)	(3.87)	(3.91)	(3.9)	(3.93)	(10.79)	(10.78)	(10.76)	(10.53)
Log book equity	0.71***	0.71***	0.71***	0.71***	0.5***	0.5***	0.5***	0.5***	0.87***	0.87***	0.87***	0.87***
	(26.19)	(25.79)	(25.69)	(26.32)	(15.29)	(15.3)	(15.23)	(15.52)	(25.05)	(24.92)	(24.82)	(25.03)
Operating profitability	0.12***	0.12***	0.13***	0.12***	0.1***	0.1***	0.1***	0.1***	0.14***	0.14***	0.14***	0.14***
	(11.21)	(11.16)	(11.28)	(11.03)	(10.98)	(11.02)	(11.34)	(10.98)	(10.05)	(10.01)	(10.05)	(9.83)
Beta	-0.07***	-0.07***	-0.07***	-0.07***	-0.06***	-0.06***	-0.05***	-0.06***	-0.06***	-0.06***	-0.06***	-0.06***
	(-5.45)	(-5.44)	(-5.22)	(-5.53)	(-5.0)	(-4.98)	(-4.75)	(-5.09)	(-4.39)	(-4.39)	(-4.24)	(-4.41)
Investment	-0.05***	-0.05***	-0.05***	-0.05***	-0.03***	-0.03***	-0.03***	-0.03***	-0.05***	-0.05***	-0.05***	-0.05***
	(-7.42)	(-7.41)	(-7.3)	(-7.17)	(-4.75)	(-4.83)	(-4.76)	(-4.67)	(-4.58)	(-4.57)	(-4.56)	(-4.53)
Dividend-to-book	0.05***	0.05***	0.06***	0.07***	0.01*	0.01*	0.02**	0.03***	0.09***	0.09***	0.09***	0.1***
	(6.51)	(6.52)	(6.67)	(8.07)	(1.95)	(2.11)	(2.44)	(3.56)	(7.41)	(7.41)	(7.44)	(8.53)
Top10 ownership		0.0	-0.0			0.02*	0.0			0.0	-0.0	
		(0.41)	(-0.29)			(2.15)	(0.41)			(0.01)	(-0.38)	
Mutual Fund ownership			0.04***				0.06***				0.03***	
			(5.47)				(6.28)				(3.96)	
DED				-0.01				-0.01				-0.0
				(-0.46)				(-0.86)				(0.0)
QIX				0.06***				0.06***				0.06***
				(5.52)				(4.78)				(5.74)
TRA				0.05***				0.05***				0.05***
				(4.57)				(4.11)				(3.66)
Panel B: Funds not owned by 401(k) pension plans												
	All Funds				**Mutual Funds**				**ETFs**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$IO^{401k}_{-i,t}(n)$	0.40***	0.41***	0.41***	0.37***	0.50***	0.50***	0.50***	0.46***	0.26**	0.27**	0.27**	0.24**
	(5.82)	(5.86)	(5.86)	(5.66)	(5.44)	(5.49)	(5.46)	(5.26)	(2.83)	(2.92)	(2.93)	(2.65)
Log market-to-book	0.53***	0.53***	0.53***	0.53***	0.31***	0.31***	0.31***	0.31***	0.74***	0.74***	0.74***	0.74***
	(9.68)	(9.64)	(9.65)	(9.6)	(5.5)	(5.48)	(5.48)	(5.55)	(9.59)	(9.55)	(9.56)	(9.58)
Log book equity	0.66***	0.66***	0.66***	0.66***	0.52***	0.52***	0.52***	0.52***	0.79***	0.79***	0.79***	0.79***
	(17.55)	(17.34)	(17.34)	(17.57)	(14.94)	(15.03)	(14.99)	(14.98)	(15.54)	(15.35)	(15.36)	(15.68)
Operating profitability	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
	(0.86)	(0.88)	(0.88)	(0.82)	(0.57)	(0.59)	(0.58)	(0.54)	(1.13)	(1.17)	(1.17)	(1.1)
Beta	-0.09***	-0.09***	-0.09***	-0.09***	-0.1***	-0.1***	-0.09***	-0.1***	-0.08***	-0.08***	-0.08***	-0.08***
	(-5.03)	(-5.12)	(-5.11)	(-5.19)	(-4.34)	(-4.44)	(-4.39)	(-4.45)	(-5.12)	(-5.16)	(-5.17)	(-5.33)
Investment	0.0	0.0	0.0	0.0	-0.0	-0.0	-0.0	-0.0	0.0	0.01	0.01	0.0
	(0.42)	(0.59)	(0.55)	(0.46)	(-0.6)	(-0.5)	(-0.62)	(-0.71)	(0.72)	(0.85)	(0.86)	(0.78)
Dividend-to-book	0.03***	0.03***	0.03***	0.03***	0.03***	0.03***	0.03***	0.03***	0.04***	0.03***	0.03***	0.04***
	(4.17)	(4.17)	(4.16)	(4.18)	(4.48)	(4.47)	(4.44)	(4.49)	(3.85)	(3.85)	(3.85)	(3.86)
Top10 ownership		-0.03**	-0.03**			-0.03	-0.03			-0.04**	-0.04**	
		(-2.24)	(-2.34)			(-1.37)	(-1.61)			(-2.87)	(-2.9)	
Mutual Fund ownership			0.01				0.03***				-0.0	
			(1.35)				(3.12)				(-0.48)	
DED				-0.02				-0.02				-0.02
				(-1.3)				(-0.92)				(-1.16)
QIX				0.03***				0.05***				0.02
				(3.32)				(3.88)				(1.63)
TRA				0.05***				0.04**				0.05***
				(3.96)				(2.68)				(3.8)

Table C.1: Demand system estimation - stock level $IO^{401k}_{t}(n)$, observations not AUM-weighted. This table reports estimates of the panel regression

$$\log\left(\frac{w_{i,t}(n)}{w_{i,t}(0)}\right) = b_0 + \beta_{0,i}\widehat{mb}_t(n) + \boldsymbol{\beta_1'}\mathbf{X}_t(n) + \beta_2\widehat{IO}^{401k}_{-i,t}(n) + \alpha_{i,t} + \tilde{u}_{i,t}(n)$$

where $\widehat{mb}_t(n)$ is the instrumented market-to-book equity, and $\mathbf{X}_t(n)$ includes the same variables as in Koijen and Yogo (2019), i.e., log book equity, operating profitability, stock market beta, investment, and dividend-to-book ratio. $IO^{401k}_{-i,t}(n)$ is the 401K plans ownership of stock n (excluding the effect through investor i), and $\alpha_{i,t}$ are fund-by-time fixed effects. The funds in the regressions are not AUM-weighted. Panel A reports results for all funds owned by 401(k) plans, while Panel B employs only those not controlling 401(k) pension assets (hence, $IO^{401k}_{-i,t}(n) \equiv IO^{401k}_{t}(n)$). Columns (1)-(3) report results using all funds (including index mutual funds and ETFs), columns (4)-(6) using only active mutual funds, and columns (7)-(9) using only active ETFs. Variables are standardized to have unit-standard deviation. Standard errors are double clustered by fund and time. The sample period is from 2007 to 2020. *p<0.1; **p<0.05; ***p<0.01.

Panel A: Mutual funds owned by pension plans									
	All Funds			Mutual Funds			ETFs		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\widehat{IO}_{-i,t}^{401k}(n)$	0.133***	0.107***	0.096***	0.179***	0.140***	0.107***	0.116***	0.093***	0.088***
	(19.89)	(25.71)	(16.79)	(10.00)	(7.62)	(5.24)	(9.63)	(8.32)	(7.36)
Log market-to-book	0.939***	0.934***	0.939***	0.441***	0.429***	0.440***	0.961***	0.953***	0.955***
	(18.47)	(18.40)	(18.62)	(5.16)	(4.78)	(4.90)	(20.88)	(21.28)	(21.49)
Log book equity	1.572***	1.580***	1.583***	0.983***	0.994***	1.002***	1.518***	1.524***	1.525***
	(25.06)	(25.08)	(25.12)	(11.86)	(11.32)	(11.40)	(63.65)	(66.93)	(66.38)
Operating profitability	0.017**	0.018**	0.019**	0.065***	0.065***	0.068***	0.019**	0.022***	0.023***
	(2.92)	(2.90)	(2.97)	(4.01)	(4.29)	(4.46)	(2.73)	(3.57)	(3.55)
Beta	-0.003	-0.006	-0.004	-0.022**	-0.026**	-0.023*	0.002	-0.002	-0.001
	(-0.91)	(-1.52)	(-1.25)	(-2.41)	(-2.53)	(-2.23)	(0.30)	(-0.25)	(-0.15)
Investment	0.010**	0.009*	0.008*	0.000	-0.005	-0.005	0.010	0.006	0.006
	(2.95)	(1.86)	(1.81)	(-0.01)	(-0.33)	(-0.37)	(1.63)	(1.02)	(0.98)
Dividend-to-book	0.003	0.005	0.004	-0.024	-0.018	-0.019	0.006	0.007	0.007
	(0.60)	(1.20)	(0.99)	(-1.7)	(-1.27)	(-1.36)	(0.69)	(0.79)	(0.76)
Top10 ownership		0.054***	0.056***		0.075***	0.077***		0.049***	0.050***
		(6.54)	(5.93)		(3.68)	(3.46)		(5.21)	(4.79)
Mutual fund ownership			0.023**			0.074***			0.013
			(2.63)			(3.73)			(1.55)
Panel B: Mutual funds not owned by pension plans									
	All Funds			Mutual Funds			ETFs		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
$\widehat{IO}_{i,t}^{401k}(n)$	0.106***	0.085***	0.080**	0.099***	0.079**	0.063*	0.097***	0.074***	0.081***
	(4.75)	(3.53)	(3.11)	(3.65)	(2.53)	(1.78)	(4.82)	(4.23)	(4.69)
Log market-to-book	0.761***	0.765***	0.765***	0.719***	0.719***	0.718***	0.808***	0.811***	0.810***
	(18.7)	(18.33)	(18.31)	(14.47)	(13.91)	(14.23)	(10.64)	(10.12)	(10.13)
Log book equity	1.340***	1.354***	1.354***	1.262***	1.264***	1.263***	1.443***	1.462***	1.462***
	(21.70)	(20.65)	(20.60)	(17.64)	(17.10)	(17.43)	(17.85)	(17.21)	(17.22)
Operating profitability	0.007	0.006	0.006	0.001	-0.001	-0.001	0.023*	0.026	0.026
	(0.58)	(0.43)	(0.45)	(0.11)	(-0.14)	(-0.07)	(1.81)	(1.61)	(1.59)
Beta	-0.010	-0.011	-0.011	-0.022	-0.027*	-0.026*	0.009	0.012*	0.012
	(-1.17)	(-1.18)	(-1.15)	(-1.66)	(-1.83)	(-1.78)	(1.35)	(1.78)	(1.68)
Investment	0.018***	0.017**	0.017**	0.016*	0.016	0.016	0.017***	0.017***	0.017***
	(3.10)	(2.89)	(2.86)	(1.84)	(1.74)	(1.70)	(3.77)	(3.15)	(3.19)
Dividend-to-book	-0.868	-0.743	-0.727	-2.162*	-2.032*	-1.971*	0.996*	1.020*	1.003
	(-1.17)	(-1.01)	(-0.98)	(-2.10)	(-2.02)	(-1.91)	(1.95)	(1.80)	(1.77)
Top10 ownership		0.036***	0.036***		0.027*	0.026*		0.045***	0.044***
		(3.26)	(3.24)		(1.87)	(1.82)		(3.57)	(3.47)
Mutual fund ownership			0.011			0.037*			-0.015*
			(1.17)			(2.21)			(-1.92)

Table C.2: Demand system estimation - stock level $IO_t^{401k}(n)$ and no TNA winsorization. This table reports estimates of the panel regression

$$\log\left(\frac{w_{i,t}(n)}{w_{i,t}(0)}\right) = b_0 + \beta_{0,i}\widehat{mb}_t(n) + \boldsymbol{\beta_1'}\mathbf{X}_t(n) + \beta_2\widehat{IO}_{-i,t}^{401k}(n) + \alpha_{i,t} + \tilde{u}_{i,t}(n)$$

where $\widehat{mb}_t(n)$ is the instrumented market-to-book equity, and $\mathbf{X}_t(n)$ includes the same variables as in Koijen and Yogo (2019), i.e., log book equity, operating profitability, stock market beta, investment, and dividend-to-book ratio. $IO_{-i,t}^{401k}(n)$ is the 401K plans ownership of stock n (excluding the effect through investor i), and $\alpha_{i,t}$ are fund-by-time fixed effects. The funds in the regressions are AUM-weighted. Panel A reports results for all funds owned by 401(k) plans, while Panel B employs only those not controlling 401(k) pension assets (hence, $IO_{-i,t}^{401k}(n) \equiv IO_t^{401k}(n)$). Columns (1)-(3) report results using all funds (including index mutual funds and ETFs), columns (4)-(6) using only active mutual funds, and columns (7)-(9) using only active ETFs. Variables are standardized to have unit-standard deviation. Standard errors are double clustered by fund and time. The sample period is from 2007 to 2020. *p<0.1; **p<0.05; ***p<0.01.

Panel A: Funds owned by 401(k) pension plans												
	All Funds				**Mutual Funds**				**ETFs**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$IO^{401k}_{-i,t-1}(n)$	0.12***	0.10***	0.09***	0.08***	0.15***	0.12***	0.09***	0.11***	0.10***	0.07***	0.07***	0.06***
	(16.09)	(14.06)	(10.56)	(7.79)	(10.86)	(9.3)	(6.88)	(7.92)	(9.47)	(7.78)	(7.04)	(4.85)
Log market-to-book	0.83***	0.83***	0.84***	0.79***	0.43***	0.43**	0.44**	0.39***	0.98***	0.97***	0.97***	0.94***
	(12.84)	(13.41)	(13.67)	(12.36)	(4.67)	(4.55)	(4.75)	(4.33)	(20.57)	(21.18)	(21.4)	(19.43)
Log book equity	1.41***	1.43***	1.43***	1.4***	0.99***	1.01***	1.02***	0.98***	1.51***	1.52***	1.52***	1.5***
	(27.22)	(26.81)	(27.01)	(27.34)	(11.91)	(11.44)	(11.58)	(11.9)	(61.12)	(64.3)	(63.55)	(60.96)
Operating profitability	0.03*	0.03*	0.03*	0.03*	0.07**	0.07**	0.07**	0.08**	0.01	0.02	0.02	0.02
	(2.49)	(2.7)	(2.75)	(2.92)	(3.76)	(3.88)	(4.06)	(4.26)	(1.56)	(1.92)	(1.93)	(2.11)
Beta	-0.0	-0.01	-0.0	-0.01	-0.02*	-0.02*	-0.02	-0.03*	0.01	0.0	0.0	-0.0
	(-0.63)	(-0.73)	(-0.5)	(-1.34)	(-2.35)	(-2.34)	(-1.98)	(-3.02)	(0.62)	(0.44)	(0.51)	(-0.04)
Investment	0.01	0.0	0.0	0.0	0.0	0.0	0.0	-0.0	0.01	0.01	0.01	0.0
	(1.41)	(0.78)	(0.79)	(0.52)	(0.31)	(0.11)	(0.14)	(-0.16)	(1.47)	(0.83)	(0.83)	(0.74)
Dividend-to-book	-0.01	-0.0	-0.0	0.01	-0.02	-0.02	-0.02	-0.01	-0.0	0.0	0.0	0.01
	(-0.89)	(-0.34)	(-0.53)	(0.78)	(-1.93)	(-1.57)	(-1.89)	(-1.04)	(-0.29)	(0.25)	(0.2)	(1.14)
Top10 ownership		0.06***	0.06**			0.06**	0.06**			0.06**	0.06**	
		(4.86)	(4.64)			(3.85)	(3.43)			(4.02)	(3.98)	
Mutual Fund ownership			0.03*				0.08***				0.01	
			(2.97)				(5.11)				(1.2)	
DED				0.04				0.04				0.05
				(1.4)				(1.13)				(1.66)
QIX				0.08***				0.09***				0.08***
				(5.23)				(4.54)				(4.88)
TRA				0.05***				0.02				0.05***
				(5.46)				(1.37)				(5.3).

Panel B: Funds not owned by 401(k) pension plans												
	All Funds				**Mutual Funds**				**ETFs**			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
$IO^{401k}_{t-1}(n)$	0.10***	0.09***	0.09***	0.07***	0.12***	0.11***	0.10***	0.09***	0.08***	0.06**	0.06**	0.04**
	(7.55)	(6.65)	(6.85)	(5.18)	(7.03)	(6.6)	(6.01)	(4.74)	(4.8)	(4.17)	(4.62)	(3.65)
Log market-to-book	0.7***	0.7***	0.7***	0.71***	0.64***	0.63***	0.63***	0.64***	0.78***	0.77***	0.77***	0.78***
	(15.46)	(14.84)	(14.81)	(15.58)	(14.67)	(15.4)	(15.46)	(14.85)	(9.75)	(9.23)	(9.21)	(9.83)
Log book equity	1.27***	1.28***	1.28***	1.29***	1.17***	1.15***	1.15***	1.18***	1.39***	1.4***	1.4***	1.39***
	(18.85)	(17.53)	(17.48)	(19.37)	(17.39)	(18.12)	(18.03)	(17.74)	(15.36)	(14.58)	(14.6)	(15.91)
Operating profitability	0.03	0.05	0.05	0.03	0.03	0.03	0.04	0.02	0.03	0.05	0.05	0.03
	(1.66)	(1.61)	(1.61)	(1.39)	(1.28)	(1.12)	(1.12)	(1.05)	(1.64)	(1.75)	(1.73)	(1.5)
Beta	-0.01	-0.01	-0.01	-0.01	-0.02	-0.03	-0.03	-0.02	0.01	0.01	0.01	0.01
	(-0.84)	(-0.79)	(-0.78)	(-1.0)	(-1.48)	(-1.62)	(-1.57)	(-1.61)	(1.23)	(1.64)	(1.53)	(1.17)
Investment	0.01	0.01	0.01	0.01	0.0	0.0	0.0	0.0	0.02	0.02	0.02	0.02*
	(1.72)	(1.53)	(1.5)	(2.06)	(0.36)	(0.24)	(0.19)	(0.67)	(2.17)	(2.02)	(2.04)	(2.35)
Dividend-to-book	-0.04	-0.04	-0.04	-0.04	-0.08**	-0.08*	-0.08**	-0.07*	-0.0	-0.01	-0.01	-0.0
	(-2.15)	(-2.1)	(-2.1)	(-1.86)	(-3.23)	(-3.24)	(-3.25)	(-2.72)	(-0.12)	(-0.67)	(-0.64)	(-0.3)
Top10 ownership		0.03*	0.03*			0.02	0.01			0.04**	0.04**	
		(2.56)	(2.54)			(1.3)	(1.15)			(3.35)	(3.36)	
Mutual Fund ownership			0.01				0.03*				-0.02*	
			(0.75)				(2.71)				(-2.37)	
DED				-0.01				-0.02				0.01
				(-0.45)				(-1.02)				(0.6)
QIX				0.07**				0.06*				0.07***
				(3.62)				(2.85)				(4.48)
TRA				0.06***				0.06*				0.06***
				(4.61)				(2.91)				(4.77)

Table C.3: Demand system estimation - lagged stock level $IO^{401k}_{t-1}(n)$. This table reports estimates of the panel regression

$$\log\left(\frac{w_{i,t}(n)}{w_{i,t}(0)}\right) = b_0 + \beta_{0,i}\widehat{mb}_t(n) + \boldsymbol{\beta_1'}\mathbf{X}_t(n) + \beta_2 IO^{401k}_{-i,t-1}(n) + \alpha_{i,t} + \tilde{u}_{i,t}(n)$$

where $\widehat{mb}_t(n)$ is the instrumented market-to-book equity, and $\mathbf{X}_t(n)$ includes the same variables as in Koijen and Yogo (2019), i.e., log book equity, operating profitability, stock market beta, investment, and dividend-to-book ratio. $IO^{401k}_{-i,t-1}(n)$ is the lagged 401K plans ownership of stock n (excluding the effect through investor i), and $\alpha_{i,t}$ are fund-by-time fixed effects. The funds in the regressions are AUM-weighted. Panel A reports results for all funds owned by 401(k) plans, while Panel B employs only those not controlling 401(k) pension assets (hence, $IO^{401k}_{-i,t}(n) \equiv IO^{401k}_{t}(n)$). Columns (1)-(3) report results using all funds (including index mutual funds and ETFs), columns (4)-(6) using only active mutual funds, and columns (7)-(9) using only active ETFs. Variables are standardized to have unit-standard deviation. Standard errors are triple clustered by fund, time and stock. The sample period is from 2007 to 2020. *p<0.1; **p<0.05; ***p<0.01.

C.4 Relevance Condition

The model in (2) is expressed in terms of the log market-to-book ratio, which we instrument as $\widehat{mb}_{-i,t}(n) = \widehat{logMB}_{-i,t}(n) = \widehat{me}_{-i,t}(n) - be_t(n)$. We test the relevance condition of our two proposed instruments by analyzing, each year, the F-statistics (Stock and Yogo, 2005; Andrews et al., 2019) obtained from the following first-stage, cross-sectional regressions of all stocks owned by each fund i:

$$\begin{aligned} IO_t^{401k}(n) = & \quad \alpha_1 + \beta_1 \widehat{IO}_{-i,t}^{401k}(n) + \beta_2 \widehat{logMB}_{-i,t}(n) + \epsilon_i(n) \quad ,\forall i \\ logMB_t(n) = & \quad \gamma_1 + \beta_3 \widehat{IO}_{-i,t}^{401k}(n) + \beta_4 \widehat{logMB}_{-i,t}(n) + \eta_i(n) \quad ,\forall i \end{aligned} \tag{C.1}$$

Specifically, each year, we first exclude funds with holdings less than 20 stocks. For the remaining funds with holdings between 20 and 200 stocks, we follow Koijen and Yogo (2019) by grouping individual funds together into one large synthetic fund, based on TNA and fund type, such that the synthetic fund owns approximately 200 stocks.[44] We then estimate (C.1), each year and for every synthetic fund separately.[45]

[44]This differs from Koijen and Yogo (2019), who set the threshold to 1,000 stocks for a given investor group in S34. We adopt a lower threshold (200) because mutual funds typically hold fewer stocks than the average institution in the S34 dataset. For instance, in our sample, the median number of stocks owned by mutual funds is 54.

[45]Differently from Koijen and Yogo (2019), we have two instruments. Hence, our $F-$tests are estimated under the null joint hypothesis $\beta_1 = \beta_2 = 0$ or $\beta_3 = \beta_4 = 0$. In other words, the alternative hypothesis is that at least one of the β coefficients is different from zero.

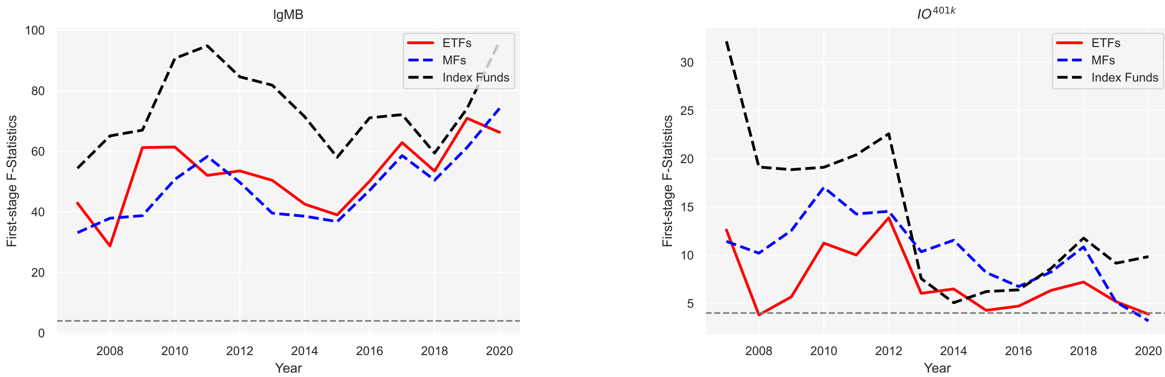


Figure C.3: ***F*-stats**.The left panel shows the 20 percentile of F-statistics for lgMB for each year and all fund types, while the right panel reports the 20 percentile of F-statistics for the stock-level $IO^{401k}_{-i,t}(n)$ for each year and all fund types. The gray dash line is the critical value for rejecting the null of weak instruments. Sample is from 2007 to 2020.

The left panel of Figure C.3 plots the 20th percentile of the F-statistic for $logMB_t(n)$ over time, while the right panel presents the same metric for stock-level $IO^{401k}(n)$, separately for active mutual funds, active ETFs, and index funds. We find that for approximately 20% of funds in our sample, $\widehat{IO}^{401k}_{-i,t}(n)$ serves as a weak instrument, with F-statistics below 4. These funds account for about 9% of the total fund-stock observations in the full sample and are excluded from our main empirical analysis. In contrast, the F-statistic for book-to-market remains consistently above the Stock and Yogo (2005) threshold of 4 for all funds. Overall, our results confirm the validity of our stock-level instruments, with only minor exceptions.

C.5 401(k) Ownership and Latent Demand

Figure C.4 reports the cross-sectional standard deviation of log latent demand for funds, weighted by assets under management. A higher standard deviation implies more extreme portfolio weights that are tilted away from observed characteristics. This figure compares the standard specification from Koijen and Yogo (2019) with two augmented versions: one that incorporates stock-level 401(k) ownership and another that additionally includes fund-level 401(k) ownership. Throughout the sample period, we find that adding stock- and fund-level 401(k) ownership reduces variation in latent demand, with both variables playing an equally important role in explaining it.

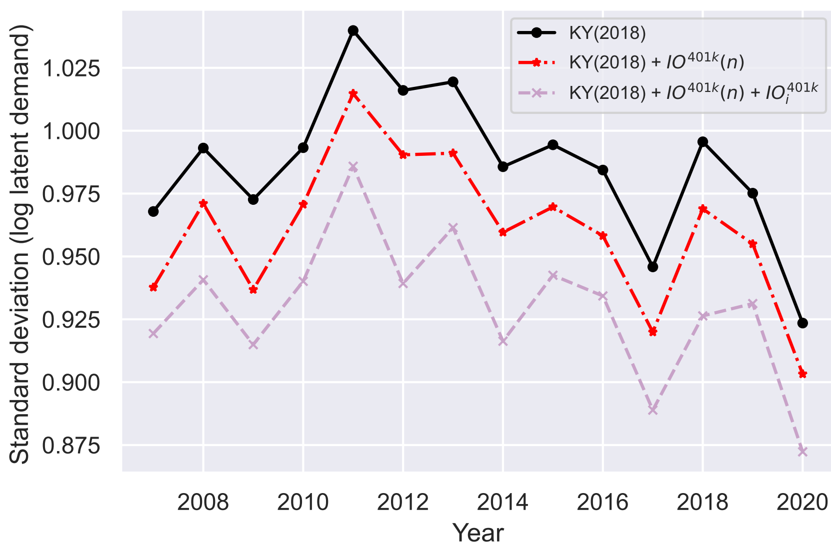


Figure C.4: Standard deviation of latent demand This figure compares the Koijen and Yogo (2019) model with versions augmented by stock-level 401(k) ownership and a further extension incorporating both stock-level and fund-level 401(k) ownership. The estimation is performed for each synthetic fund at the end of each year. This figure presents the standard deviation of the log latent demand for funds that manage 401(k) assets, weighted by their AUM. The sample period spans from 2007 to 2020.

Internet Appendix D Pseudo Algorithm for the Counterfactual

In this Appendix, we provide a pseudo algorithm for the counterfactual design.

1. For each pasive investor $j \in \mathcal{P}$, estimate the constant $\widehat{\text{const}_j}$ from the equation:

$$\frac{w_{j,t}(n)}{w_j(0)} = \exp\left\{me(n) + \text{const}_j\right\}$$

2. Compute Adjusted AUM for Active Funds: For each active fund $k \in \mathcal{A}$, calculate the AUM excluding assets from each of the 401(k) plan p:

$$\widetilde{AUM}_{k,t} = AUM_{k,t} - \sum_p AUM_{p,k,t} \qquad k \in \mathcal{A}$$

This represents the expected new AUM for active funds offered by pension plans. Of course, in equilibrium, since the price of stocks will change, the AUM will do as well.

3. Reallocate 401(k) Assets to Passive Funds: Redistribute the assets of each pension plan p from active funds (which are not offered anymore), $\sum_{k \in \mathcal{A}} AUM_{p,k,t}$, to passive funds currently offered by the plan $j \in \mathcal{P}(p)$:

$$F_{p,j,t} = \frac{AUM_{j,t}}{\sum_{i \in \mathcal{P}(p)} AUM_{i,t}} \times \sum_{k \in \mathcal{A}(p)} AUM_{p,k,t} \qquad j \in \mathcal{P}(p)$$

4. Solve for Counterfactual Prices: Determine the counterfactual prices $P_t^{CF}(n)$ that clear the market by iterating on Equation (13) in the text until convergence, using the algorithm in Koijen and Yogo (2019, Appendix C). As discussed in Section 4.2.1 a key simplification arises from the independence of passive fund weights from stock-level 401(k) ownership. In this case equation (13) becomes:

$$P_t^{CF}(n) = \sum_{i=\mathcal{A},\mathcal{P},1\ldots,I} AUM_{i,t}^{CF}\left(\mathbf{P}_t^{CF}\right) w_{i,t}^{CF}(n; \mathbf{P}_t^{CF})$$

where the counterfactual AUM distribution, AUM^{CF}, is implied from equation (12) at the converged vector of equity prices. Also, note that the equilibrium condition includes active mutual fund investors since they still have positive, albeit lower, AUM (see step 2).

5. Calculate Counterfactual 401(k) Stock-Level Characteristic: The weights $w_{i,t}^{CF}$ for all $i \neq \mathcal{P}$ depend on the counterfactual stock-level 401(k) characteristic:

$$IO_t^{401k,CF}(n) = \frac{\sum_{j\in\mathcal{P}} \left(\sum_{p=1}^{M} AUM_{p,j,t}^{CF}\right) \times w_{j,t}^{CF}(n)}{P_t^{CF}(n)}$$

where

$$\frac{w_{j,t}^{CF}(n)}{w_j(0)} = \exp\left\{me^{CF}(n) + \widehat{\text{const}_j}\right\} \qquad j \in \mathcal{P}$$

and the $\widehat{\text{const}_j}$ is from step 1, as demand coefficients are held constant during the counterfactual scenario.

We generate an average of 200 synthetic funds per year, while for ETFs, we create approximately 60 groups annually. Each synthetic fund has its own coefficient on stock-level 401(k) ownership. The criteria used align with the two-stage regression approach: for each year, we first exclude funds holding fewer than 20 stocks. Among the remaining funds with fewer than 200 holdings, we apply the method of Koijen and Yogo (2019), grouping them into synthetic funds based on total net assets (TNA) and fund type to ensure that each synthetic fund's holdings are close to 200 stocks.[46]

D.1 Additional Results

The NEF measure defined in Equation (16) relies on observed quantities, such as $AUM_{p,j,t}^{NEW}$, rather than counterfactual values like $AUM_{p,j,t}^{CF}$ derived from the market-clearing condition (13). As such, it serves as a gauge of net expected flows, which may differ from those obtained in the counterfactual equilibrium. Importantly, Figure D.1 demonstrates that our NEF provides a reliable proxy for actual realized flows.

[46]This is done for both our scenarios where only funds and all investors care about the 401(k) characteristic.

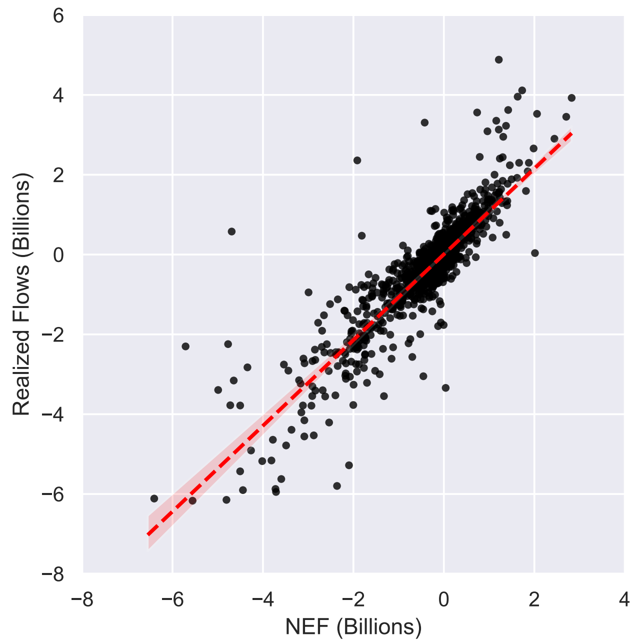


Figure D.1: Counterfactual: NEF versus Realized Flows. This figure plots the realized flows, which depends on counterfactual values of $AUM^{CF}_{p,j,t}$ derived from the market-clearing condition (13), against the net expected flow measure $NEF^{401k}(n)$ defined in equation (16).

Figure D.2 mirrors Figure 6, but while the latter considers a scenario where only funds tilt toward stock-level 401(k) ownership, Figure D.2 examines a broader case where all investors, except households, exhibit this preference. The patterns remain largely consistent with those discussed in Section 4.3, with the key difference being the magnitude of the effects. In this expanded scenario, firms are more likely to face representative owners with stronger preferences for stock-level 401(k) ownership, leading to more pronounced repricing.

Finally, Figure D.3 serves as the counterpart to Figure 7, bottom panel, representing a scenario where investors do not account for stock-level 401(k) ownership. This figure acts as a placebo test. As expected, the relationship between counterfactual returns and stock-level 401(k) ownership is now a flat line, since $\beta_{401k} = 0$ for all investors. Consequently, changes in stock-level 401(k) ownership have no impact on counterfactual prices.

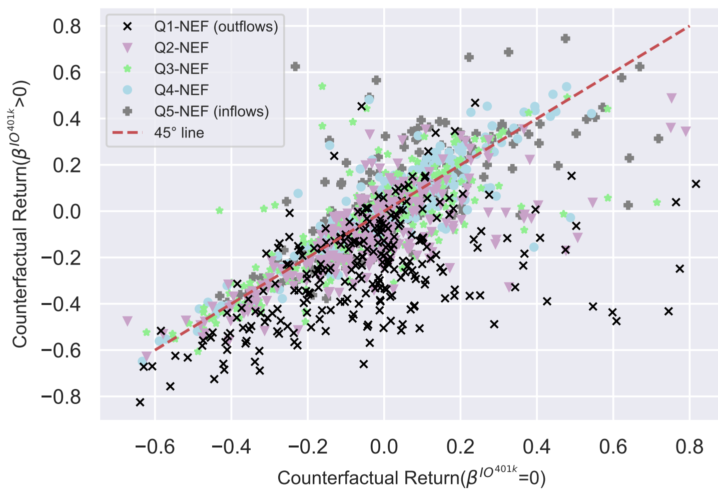


Figure D.2: Counterfactual returns: Impact of All Investors' 401(k) Ownership Preferences. This figure presents a scatter plot of the counterfactual stock returns, comparing the scenario where all investors, except households, display a preference for 401(k) stock-level ownership ($\beta^{IO^{401k}} > 0$, y-axis) to the scenario where all investors are indifferent to this characteristic ($\beta_i^{IO^{401k}} = 0$ for all investors i, x-axis). Stocks are color-coded based on their net expected flow measure, $NEF^{401k}(n)$ (%).

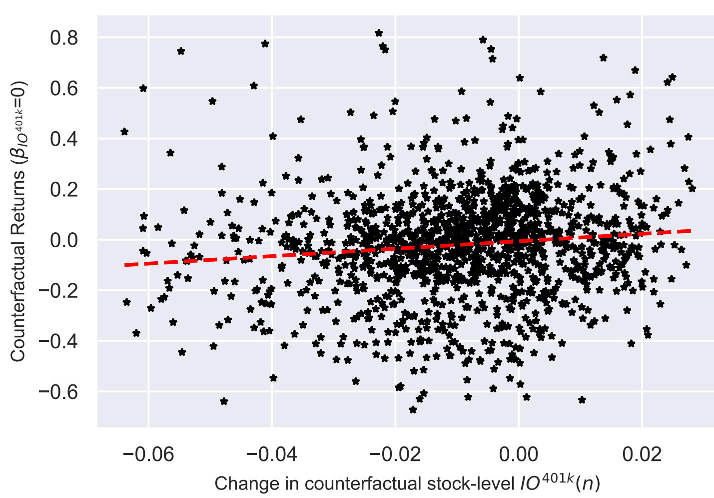


Figure D.3: Counterfactual Returns vs. Stock-Level 401(k) Ownership: Placebo Test Without Investor Preferences for 401(k) Ownership. This figure presents a scatter plot of counterfactual stock returns, measured as the difference between the log of the counterfactual price (assuming no investor preferences for stock-level 401(k) ownership, i.e., $\beta_{IO^{401k,funds}} = 0$) and the log of the observed price before the reallocation of AUM from active to passive funds. The x-axis represents the change in stock-level 401(k) ownership.

Internet Appendix E Equilibrium Price Impact of 401(k) Plans

In this section we quantify the equilibrium price impact of a change in 401(k) stock-level ownership for firm n, accounting for the trading of *all* investors. Specifically, we estimate

$$\frac{\partial p_t(n)}{\partial IO_t^{401k}(n)} \tag{E.2}$$

where p is the log price of stock n. Following Koijen and Yogo (2019) and Noh and Oh (2020), this derivative can be computed analytically, at any time t, as the diagonal elements of the matrix $\mathbf{M}$:[47]

$$\mathbf{M} = \left(\mathbf{I} - \sum_i \beta_{0,i} A_i \mathbf{H}^{-1}\mathbf{G_i}\right)^{-1} \left(\sum_i \beta_{2,i} A_i \mathbf{H}^{-1}\mathbf{G_i}\right), \tag{E.3}$$

where we recall that $\beta_{0,i}$ is the loading of investor i on market-to-book, and $\beta_{2,i}$ is the coefficient on 401(k) ownership (see equation (2)). The matrices $\mathbf{H} = \sum_{i=1}^{I} A_i \mathrm{diag}(\mathbf{w}_i)$ and $\mathbf{G}_i = \mathrm{diag}(\mathbf{w}_i) - \mathbf{w}_i\mathbf{w}_i'$ instead do not depend on estimated parameters, but only on investors' weights $\mathbf{w}$. Finally, A_i denotes the assets under management of investor i.

The n-th diagonal entry of $\mathbf{M}$, $M_{n,n}$, captures two effects. First, the matrix inside the inverse in equation (E.3) is the aggregate demand elasticity (Koijen and Yogo, 2019), and its diagonal elements are strictly positive when $\beta_{0,i} < 1$ for all investors. If a firm is held by less price elastic investors, then the firm price will react more due to institutional demand for the $IO_t^{401k}(n)$ characteristic. Second, the n-th diagonal entry of the matrix outside the inverse can be written as $\frac{\sum_i \beta_{2,i} A_i w_i(n)(1-w_i(n))}{\sum_i A_i w_i(n)}$, and represents an AUM weighted average of the coefficients on the 401(k) stock-level ownerhship (multiplied by $1 - w_i(n)$). This implies that the price pressure is larger if a firm faces owners that are large and exhibit a high coefficient on the $IO_t^{401k}(n)$. In other words, the institutional price pressure that a given firm n receives due to a change in the level of 401(k) ownership is a weighted average of $IO_t^{401k}(n)$ coefficients of its institutional owners, adjusted for their demand

[47] To compute this expression one has to exploit the identity $\mathbf{p} = \log\left(\sum_i A_i \mathbf{w}_i\right) - \mathbf{s}$ (where $\mathbf{s}$ denotes the vector of shares outstanding) which holds by market clearing. See Appendix A in Noh and Oh (2020) for additional details.

elasticities.

To compute the price impact $M_{n,n}$ we need to consider the entire investor universe, i.e., not only mutual funds and ETFs. To this end, we use data on institutional common stock holdings from the Thomson Reuters Institutional Holdings Database (s34 file). We follow the Koijen and Yogo (2019) classification of institutions into six types (i.e., $i = 1, \ldots, 6$): banks, insurance companies, investment advisors, mutual funds, pension funds, and other 13F institutions. We recall that the s34 file provides a different level of granularity relative to our analysis in Section 4, since it reports aggregate holdings at the investor level (e.g., for all funds managed by Fidelity).[48]

Figure E.4 displays the two key ingredients required to compute the price impact: the coefficient on market-to-book driving demand elasticities (Panel A) and the coefficient on 401(k) stock-level ownership (Panel B) for each of the six groups of investors. These coefficients are estimated year-by-year by GMM, accounting for zero holdings, from model (2), under moment condition (3).[49] We confirm the results in Koijen and Yogo (2019) that mutual funds have less elastic demand than investment advisors for most of our sample period, and that insurance companies and pension funds have become less elastic over time. The coefficient in Panel B captures institutional demand for 401(k) stock-level ownership. When positive, it implies that investor i allocates at time t more weight to stocks with higher 401(k) ownership, controlling for other stock characteristics. We see that mutual funds, banks and insurance companies tilt their portfolio toward stocks with high-level of 401(k) ownership more than other types of institutions. In contrast, investment advisors do not manifest such a tilt. Interestingly, the tilt of pension funds toward stocks with high level of $IO_t^{401k}(n)$ increases over our sample period suggesting an intricate relation between the sample of funds offered by 401(k) plans, their holdings, 401(k) plan investor preferences, and the type of individual stocks preferred by pension plans

[48]Mindful of potential gaps in coverage of institutional holdings in the s34 files, we validate our price impact results by replacing s34 data with data on 13F filings from Backus et al. (2021) in Internet Appendix F.2.

[49]To obtain the price impact, we estimate the coefficients as in Koijen and Yogo (2019). For institutions with more than 1,000 stocks in their holdings, we estimate coefficients by institution. For the remaining institutions, we group them by type (e.g., mutual funds) such that on average each group holds 2,000 stocks at any point in time. Variables are standardized within each institution (or group) and for each year. We instrument market-to-book with $z_{i,t}(n)$ as usual.

(e.g., green stocks). Finally, the evidence in Panel B for investors other than mutual funds emphasizes the relevance of stock-level 401(k) ownership as an important characteristic while further alleviating endogeneity concerns: we use stock holdings of banks, insurance, etc., as left hand side variables, while we employ only mutual funds and ETFs holdings in the construction of our stock-level 401k ownership (right hand side variable).

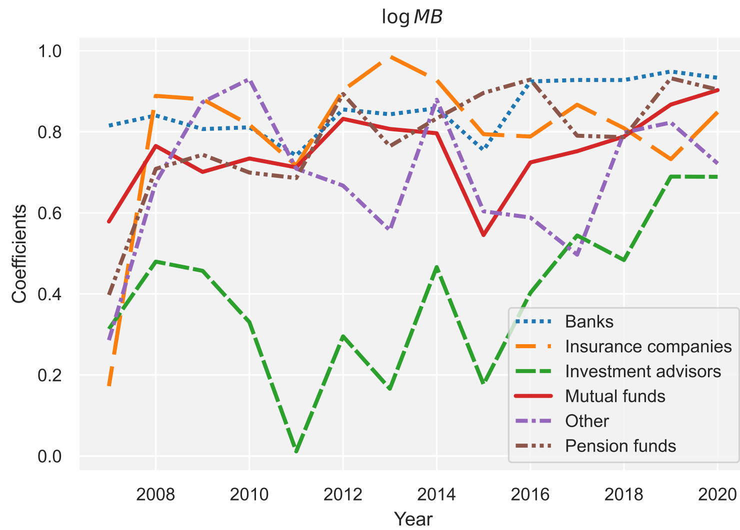
Given estimates of $\beta_{0,i,t}$ and $\beta_{2,i,t}$ for each investor, we can calculate, each time period t, the firm-level institutional pressure with respect to 401(k) ownership. The top left panel in Figure E.5 shows the cross-sectional distribution of price impact across all stocks. The aggregate price impact for the median stock (solid black line) has generally increased over time, and the cross-sectional spread has also significantly expanded over our sample period. The stronger effect over time can be related to the shift from active to passive investing of the last decade, since equation (E.3) implies that the presence of more inelastic investors results in larger price pressure. A one standard deviation increase in 401(k) ownership leads to a price impact (for the median stock) slightly less than 20 percent in 2007 and of about 60 percent in 2020.[50] The remaining panels display the aggregate price impact for extreme quintile portfolios of stocks sorted on book-to-market (top right panel), market beta and size (bottom left and right panels, respectively). We observe that the average price impact has increased for large stocks with a sharp jump in 2015, while it has remained relatively stable for small stocks. This resonates well with Haddad et al. (2025) who find that investor elasticities are lower for larger stocks (i.e., investors are more reluctant to change their positions for large stocks than for small stocks), given tracking error concerns.[51]

We do not observe noticeable differences for stocks sorted on book-to-market or betas, which suggests that a change in 401(k) stock-level ownership variable has the same price impact on growth and value stocks. For book-to-market and betas-sorted portfolios, we again observe a positive low-frequency trend of price impact from 2008 to 2020. However, we also observe an interesting cyclical pattern around this trend, particularly for value

[50] The standard deviation of 401(k) ownership is 1% in 2007 and 2% in 2020; thus, the price of the median stock increases by 0.2% and 1.2%, respectively.

[51] In the U.S. stock market, large corporations like Apple make up a substantial fraction of total market capitalization and, as a consequence, a large change in those portfolio weights would cause a substantial impact on an institution's total portfolio return.

Panel A: Log Market-to-Book



Panel B: Stock-level 401(k) Ownership

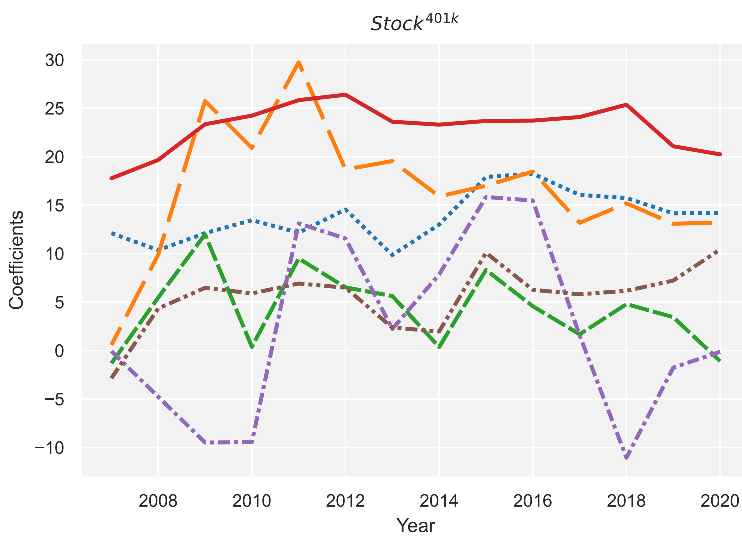


Figure E.4: Price impact: relevant coefficients. This figure shows the annual coefficient on log market-to-book (top panel) and stock-level 401(k) ownership (bottom panel) for financial institutions in Thomson Reuters holding (s34) estimated annually by GMM with zero weights. Variables are standardized (within each year) to make coefficients comparable. We report the cross-sectional mean of the estimated coefficients by institution type, weighted by assets under management. The coefficient on 401(k) ownership is multiplied by 100. The sample period is from 2007 through 2020.

and high-beta stocks.

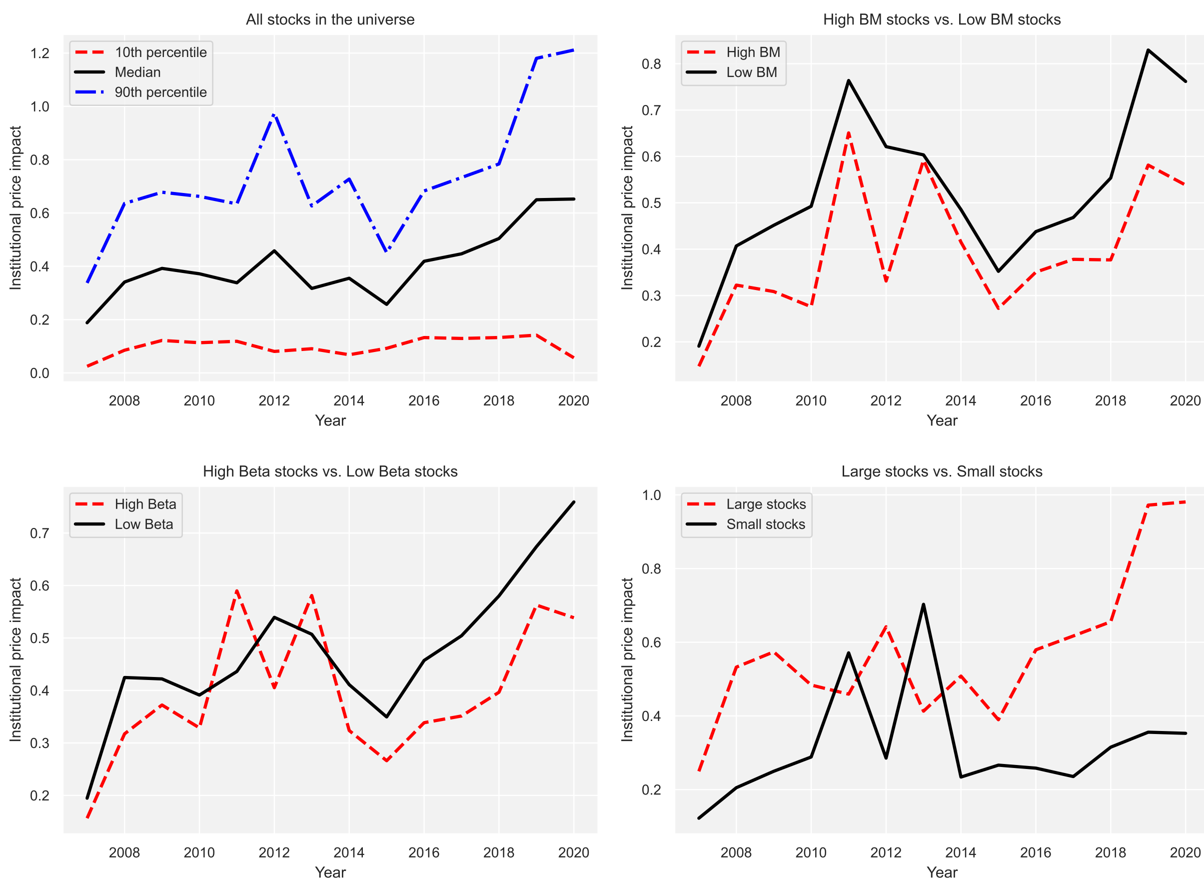


Figure E.5: Institutional price impact. This figure shows the price impact of a change in the stock-level 401(k) ownership estimated through the diagonal elements of the matrix **M** defined in (E.3). The top left panel shows the 10th percentile, median and 90th percentile of price impact across all stocks. The remaining panels plot the average price impact across the stocks in the top quintile and bottom quintile of portfolios sorted on (top right), beta (bottom left), and size (bottom right), using NYSE break points as cutoffs.

Internet Appendix F Holdings Data: Additional Analysis and Robustness

F.1 Thomson Reuters s34 Holdings

The analysis in Section 4 relies on data from Morningstar, which provides detailed holdings of *individual* mutual funds and ETFs. Instead, the analysis in Section Internet Appendix E relies on the Thomson Reuters' s34 file, which provides aggregated holdings of *all funds* under the manager's control (Koijen and Yogo (2019), Koijen et al. (2024)).

Table F.1 repeats the same analysis presented in Table 2 but at the fund family level, i.e., using the s34 data. Importantly, the coefficient on stock-level ownership remains large and significant. In particular, we find that the coefficient of .218 is close in magnitude to the one reported in column (4) of Table 2, despite the fact that s34 fund-family holdings blend together ETFs and mutual funds.

Thomson Reuters (s34) holdings			
	Coefficient	s.e.	t-stat
$IO_t^{401k}(n)$	0.218***	0.037	5.830
Log market-to-book	1.514***	0.162	9.350
Log book equity	1.893***	0.067	28.300
Operating profitability	0.006	0.007	0.830
Beta	0.056*	0.030	1.840
Investment	0.036*	0.020	1.810
Dividend-to-book	-0.135***	0.031	-4.300

Table F.1: Demand system estimation - Stock level $IO_t^{401k}(n)$ with s34 holdings. This table reports estimates of the panel regression

$$\log\left(\frac{w_{i,t}(n)}{w_{i,t}(0)}\right) = b_0 + \beta_{0,i}\widehat{mb}_t(n) + \boldsymbol{\beta_1'}\mathbf{X}_t(n) + \beta_2 IO_t^{401k}(n) + \alpha_{i,t} + \tilde{u}_{i,t}(n)$$

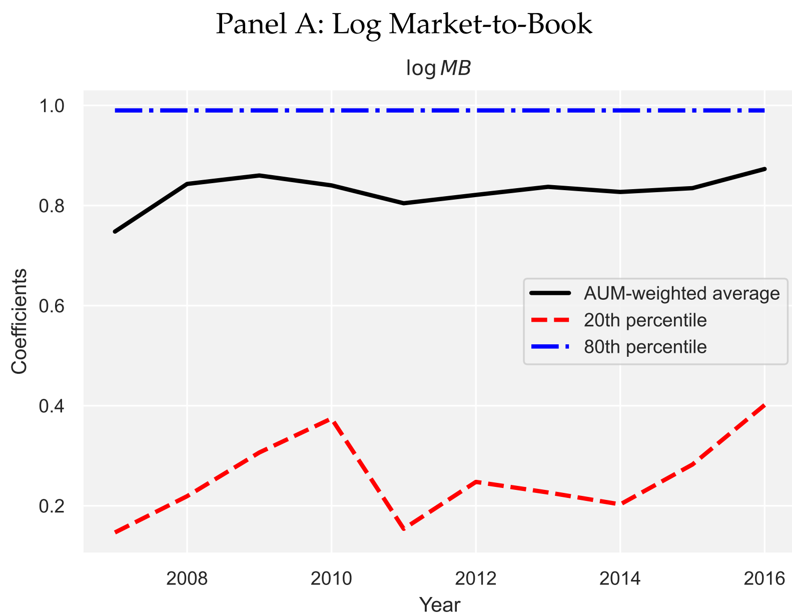
where $\widehat{mb}_t(n)$ is the instrumented market-to-book equity, and $\mathbf{X}_t(n)$ includes the same variables as in Koijen and Yogo (2019), e.g., , log book equity, operating profitability, stock market beta, investment, and dividend-to-book ratio. $IO_t^{401k}(n)$ is the 401K plans ownership of stock n, and $\alpha_{i,t}$ are manager-by-year fixed effects. The mutual fund institutions in the regressions are AUM-weighted. Variables are standardized. Standard errors are double clustered by fund institution and time. The sample period is from 2007 to 2020. *p<0.1; **p<0.05; ***p<0.01.

The coefficients on the other characteristics, e.g., beta, investment, and dividend-to-book are small in both datasets. Overall, it appears that the empirical results using the s34 dataset are in line with those reported in Section 4 and, thus, the analysis in Section Internet Appendix E is informative of the equilibrium price impact of a change in 401(k) stock-level ownerhship.

F.2 Holdings scraped directly from 13F filings

In this section, we repeat our computation of the equilibrium price impact presented in Section Internet Appendix E using the 13F filings data provided by Backus et al. (2021). These authors collected 13F filings from the SEC's EDGAR database since electronic filing was made mandatory in 1999, and addressed gaps in coverage and errors that appear in commercial datasets of institutional holdings (e.g., Thomson Reuters). The disadvantage of such dataset is that we cannot anymore exploit the Koijen and Yogo (2019) classification of institutions into six types. Thus, in the estimation, we abstract from investor types and (1) keep institutions with more than 1,000 strictly positive holdings separate; (2) group institutions with fewer than 1,000 holdings based on TNA, so that each group has on average 2,000 holdings.

Figure F.1 is the counterpart of Figure E.4. Importantly, both the coefficient governing the elasticity of demand, and the coefficient on 401(k) stock-level ownership display a similar range in terms of magnitude across the two datasets. It is therefore not surprising that the cross-sectional distribution of aggregate price impact across stocks reported in the top left panel of Figure F.2 remains economically sizable: a one standard deviation increase in 401(k) ownership, around 1.3% in 2007 and 1.6% in 2016, leads to a price impact (for the median stock) slightly less than 40 percent in 2007 and about 90 percent in 2016. Similarly to the s34 dataset, we observe a stronger price impact for large stocks (bottom right panel), with a sharp increase in 2015, and little difference for stocks sorted on market betas (bottom left panel). The main difference across the two datasets is observed for stocks sorted on book-to-market. In particular, the scraped data of Backus et al. (2021) suggest a larger impact for growth stocks.



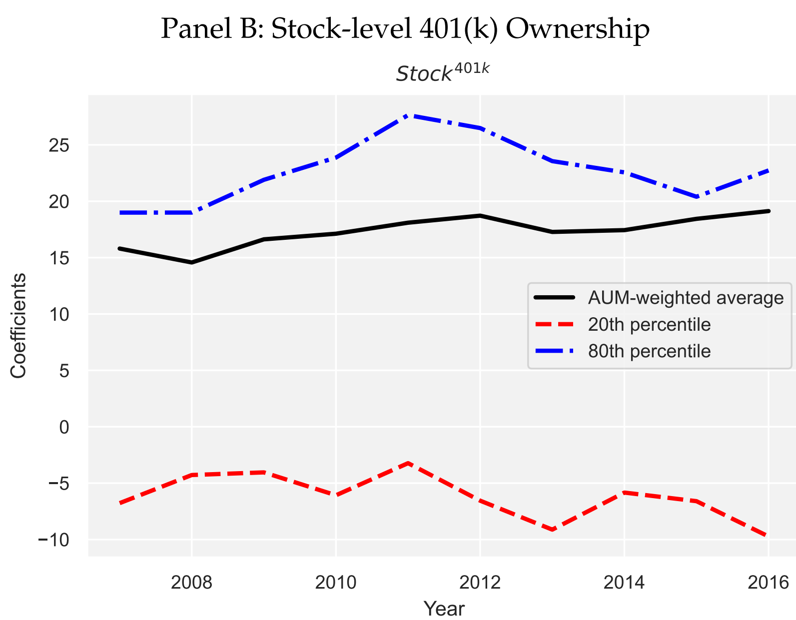


Figure F.1: Price impact: relevant coefficients. This figure shows the annual coefficient on log market-to-book (top panel) and stock-level 401(k) ownership (bottom panel) for financial institutions in Backus, Conlon and Sinkinson (2021), estimated annually, by GMM with zero weights. Variables are standardized (within each year) to make coefficients comparable. We report the cross-sectional mean of the estimated coefficients by institutions, weighted by assets under management. The coefficient on 401(k) ownership is multiplied by 100. The sample period is from 2007 through 2016.

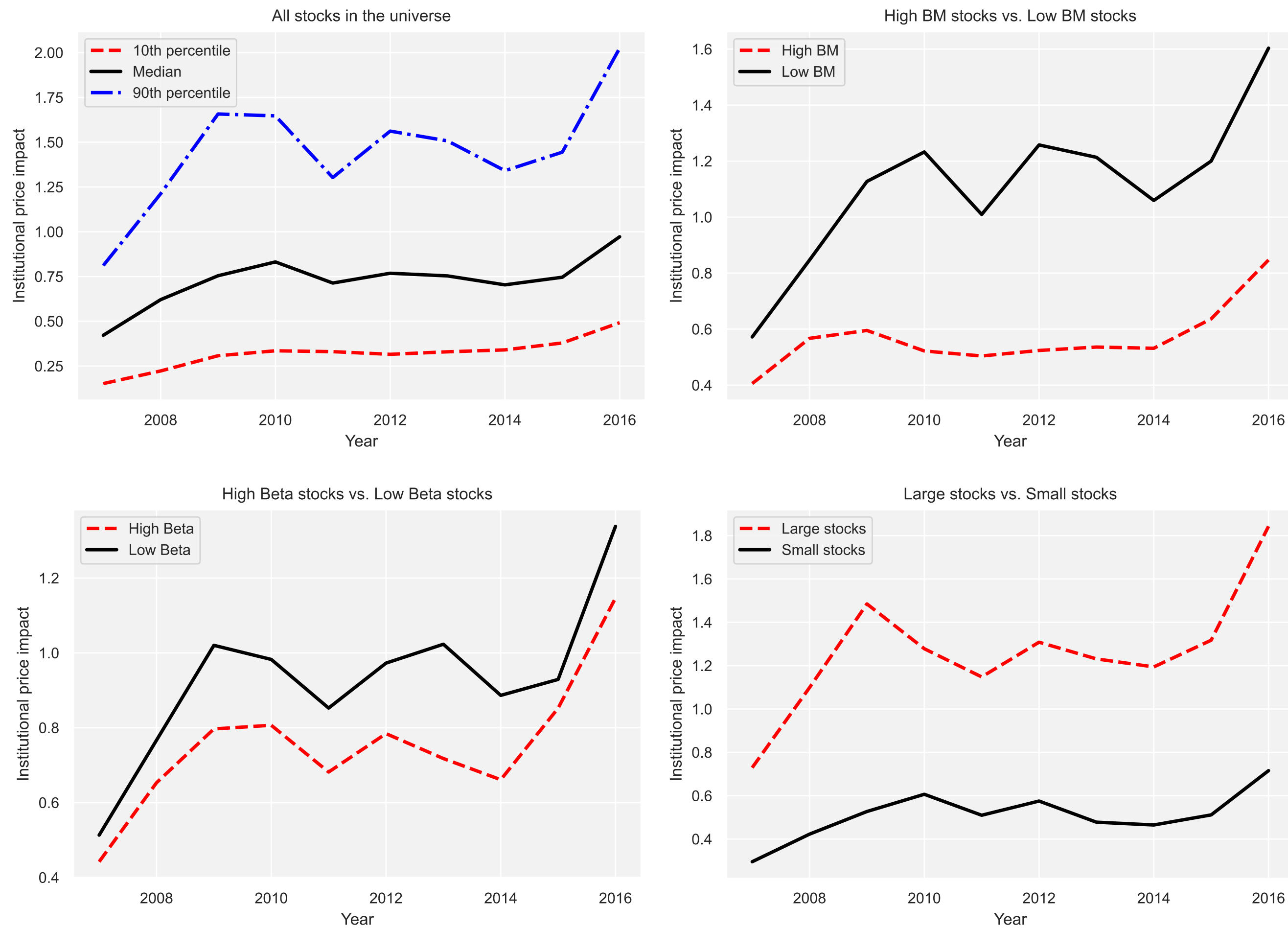


Figure F.2: Institutional price impact. This figure shows the price impact of a change in the stock-level 401(k) ownership estimated through the diagonal elements of the matrix M defined in (E.3) using holdings data from Backus, Conlon and Sinkinson (2021). The top left panel shows the 10th percentile, median and 90th percentile of price impact across all stocks. The remaining panels plot the average price impact across the stocks in the top quintile and bottom quintile of portfolios sorted on (top right), beta (bottom left), and size (bottom right), using NYSE break points as cutoffs.